

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS QUÍMICAS
Departamento de Bioquímica y Biología Molecular I



TESIS DOCTORAL

**Estudio estructural y funcional del sistema de transferencia
de azufre CSD ("Cysteine Sulfinatase Desulfinate") de "E. coli"
y proteínas con las que interacciona**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Miguel López Estepa

Directora

M^a Cristina Vega Fernández

Madrid, 2018

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS QUÍMICAS
Departamento de Bioquímica y Biología Molecular I



Tesis doctoral

Estudio estructural y funcional del sistema
de transferencia de azufre CSD (*Cysteine Sulfinat*
Desulfinate) de *E. coli* y proteínas con las que interacciona

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
presentada por

Miguel López Estepa
Directora de tesis
M.^a Cristina Vega Fernández



Centro de Investigaciones Biológicas
Madrid, 2017

AGRADECIMIENTOS

No cabe duda de que llegar hasta aquí es consecuencia de un esfuerzo y dedicación personales, pero nada de esto puede hacerse realidad sin el apoyo y paciencia de aquellos que están en el día a día. Por ello quiero agradecer que esta tesis se haya hecho realidad a diferentes personas.

En primer lugar y de manera muy especial a mi padres, Mayte y Miguel Ángel, y mi hermano, Enrique, por su paciencia diaria y constante, porque sé que no ha sido ni es fácil. Y por que sin ellos no sería el hombre y persona que soy hoy.

En segundo lugar a mi directora de tesis, Cristina, y a Francisco por depositar su confianza en mí desde el principio, cuando acababa de licenciarme y no tenía claro cual podría ser el primer paso en mi carrera. Otros proyectos te forman de manera personal y profesional, pero la realización de una tesis doctoral es sin duda un reto especial y ellos me encaminaron hacia este viaje.

También me acuerdo de toda mi familia, la que ya no está y la que acaba de llegar, el pequeño Beltrán, porque siempre han estado atentos de mí en todo momento.

No me puedo olvidar de mis amigos de toda la vida que siempre han estado para lo que sea y se necesite. Y en especial a Damián y Laura, dos grandes amigos y personas muy especiales para mí con las que he vivido grandes momentos.

Por último a mis compañeros de tesis, Alejandra, Esther, Virginia, Juan y Sara, los cuales me han apoyado también para la realización de esta tesis. Y en especial a Javi, que ha sido compañero en esta tesis pero también amigo.

Nada de esto hubiese sido posible sin vosotros y muchos otros que de manera directa o indirecta me habéis apoyado día a día. Muchas gracias a todos.

ÍNDICE

Resumen	17
Summary	23
1. Introducción	31
1.1. Tráfico de azufre en <i>E. coli</i>	31
1.1.1. Tipo de <i>clusters</i> [Fe-S]	31
1.1.2. Modificaciones de nucleósidos, vitaminas y cofactores	33
1.2. Cisteína desulfurasas: clasificación	35
1.2.1. Cofactor PLP: un mismo cofactor para una miríada de reacciones	36
1.2.2. Familias de enzimas dependientes de PLP	38
1.2.3. Mecanismo catalítico de las cisteína desulfurasas	39
1.2.4. Reacción de transpersulfuración	42
1.3. Sistemas enzimáticos bacterianos relacionados con el tráfico de azufre	42
1.3.1. Sistema Isc (<i>Iron sulfur cluster</i>)	42
1.3.2. Sistema Nif (<i>Nitrogen fixation</i>)	43
1.3.3. Sistema Suf (<i>Sulfur formation</i>)	44
1.3.4. Sistema CSD (<i>Cysteine Sulfinate Desulfinase</i>)	46
1.4. Miembros del sistema CSD	46
1.4.1. CsdA, cisteína desulfurasa	47
1.4.2. CsdE, aceptora de átomos de azufre de CsdA	48
1.4.3. TcdA, proteína modificadora de ciertos tipos de ARNt	51
1.4.3.1. Homólogos de TcdA en eucariotas	52
1.4.3.2. Homólogos de TcdA en procariotas	53
1.4.3.3. Adenilación	56
1.4.3.4. Actividad TcdA	57
1.4.3.5. Modificación sobre moléculas de ARNt	58
1.4.3.6. Modificaciones t ⁶ A y ct ⁶ A en la posición 37.....	60
1.5. Interactores de CsdE	61
1.5.1. Fosforribosil transferasa o ribosa fosfato pirofosfatasa (PrsA)	62
1.5.2. Actividad PRPP sintetasa	63
1.5.3. Clasificación de PRPP sintetasas	64
1.5.4. Estructura cuaternaria de las enzimas PRPP sintetasas	65
1.5.4.1. La enzima PRPP sintetasa como diana terapéutica	65
1.5.4.2. Estructuras resueltas de PRPP sintetasas	66
1.5.5. Estudio de OPRT y su relación con la PRPP sintetasa: entrada en la ruta de síntesis <i>de novo</i> de pirimidinas	67
1.5.5.1. La enzima OPRT como diana terapéutica	70
1.5.5.2. Estructuras resueltas de OPRT	71
2. Objetivos	75
3. Materiales y Métodos	79
3.1. Herramientas bioinformáticas	79
3.1.1. Secuencias de ADN y proteínas	79
3.1.2. Estructura 3D: PDB (<i>Protein Data Bank</i>)	79
3.2. Métodos generales en clonaje, expresión y purificación de proteínas recombinantes de <i>E. coli</i>	80
3.2.1. Clonaje	80
3.2.1.1. PCR (<i>Polymerase Chain Reaction</i>)	80
3.2.1.2. Electroforesis de ADN	81

3.2.1.3.	Digestión y ligación	82
3.2.1.4.	Transformación mediante choque térmico	82
3.2.1.5.	Mutagénesis dirigida	83
3.2.1.6.	Pruebas de expresión a pequeña escala y posterior escalado	86
3.2.1.7.	Producción y obtención de ARN de transferencia	87
3.2.1.8.	Electroforesis de proteínas en condiciones desnaturalizantes (PAGE-SDS) y nativas	88
3.2.2.	Purificación de proteínas	88
3.2.2.1.	Lisis celular	88
3.2.2.2.	Medición de concentración de proteína y ácido nucleicos	89
3.2.2.3.	Purificación por cola de afinidad	89
3.2.2.4.	Intercambio iónico	90
3.2.2.5.	Filtración en gel	91
3.3.	Métodos de análisis	92
3.3.1.	<i>Cross-linking</i>	92
3.3.2.	DLS (<i>Dynamic Light Scattering</i>)	93
3.3.3.	SPR (<i>Surface Plasmon Resonance</i>) o resonancia de plasmón de superficie	94
3.3.4.	Espectrometría de masas	95
3.3.5.	Actividad enzimática de OPRT	96
3.4.	Métodos de caracterización estructural	96
3.4.1.	AUC, ultracentrífuga analítica	96
3.4.2.	Resonancia magnética nuclear (RMN)	97
3.4.3.	SAXS (<i>Small Angle X-ray Scattering</i>)	99
3.5.	Métodos de cristalografía de rayos-X	101
3.5.1.	Métodos de cristalización	101
3.5.2.	Recogida y procesamiento de datos	106
3.5.3.	Problema de las fases	111
3.5.3.1.	Reemplazamiento Molecular (MR)	112
3.5.3.2.	SAD (<i>Single-Wavelength Anomalous Dispersion</i>)	114
3.5.3.3.	Refinamiento y validación	115
4.	Resultados	121
4.1.	Cisteína desulfurasa del sistema CSD (CsdA)	121
4.1.1.	Clonaje de CsdA	121
4.1.2.	Expresión de CsdA	121
4.1.3.	Purificación de CsdA	122
4.1.4.	Clonaje de los mutantes de CsdA (H120A y K250A)	125
4.1.5.	Expresión y purificación de los mutantes de CsdA (H120A y K250A)	127
4.1.6.	Persulfuración de CsdA	131
4.1.7.	Cristalización de CsdA nativa, persulfurada y oxidada	132
4.1.8.	Colección y procesamiento de datos de difracción de rayos-X de la proteína nativa persulfurada y S-sulfatada	133
4.1.9.	Resolución estructural y refinamiento de CsdA nativa persulfurada, oxidada	135
4.1.10.	CsdA, la cisteína desulfurasa de <i>E. coli</i> , presenta tres estados de oxidación de la cisteína catalítica (apo, persulfuración y S-sulfatada)	137
4.1.11.	Cristalización de los mutantes CsdA H120A y CsdA K250A	139
4.1.12.	Colección y procesamiento de datos de difracción de rayos-X de los mutantes de CsdA	139
4.1.13.	Resolución estructural y refinamiento de los mutantes de CsdA	140
4.1.14.	Los mutantes de CsdA estudiados (H120A y K250A) no muestran cambios significativos en la estructura respecto a la proteína nativa	142
4.2.	CsdE	143
4.2.1.	Clonaje de CsdE	143

4.2.2.	Expresión de CsdE	144
4.2.3.	Purificación de CsdE	144
4.2.4.	Clonaje de mutantes de CsdE	147
4.2.5.	Expresión y purificación de mutantes de CsdE	147
4.3.	Complejo de CsdA-CsdE	150
4.3.1.	Ensamblado y purificación de CsdA-CsdE	150
4.3.2.	Persulfuración del complejo CsdA-CsdE	150
4.3.3.	Caracterización de la interacción CsdA-CsdE mediante SPR.....	151
4.3.4.	Cambio conformacional entre CsdE libre (RMN, pdb 1ni7) y CsdE unida a CsdA ..	153
4.3.5.	SAXS del complejo CsdA-CsdE	155
4.3.6.	Cristalización del complejo (CsdAper-CsdEper) ₂	157
4.3.7.	Colección y procesamiento de datos de difracción de rayos-X	157
4.3.8.	Resolución estructural y refinamiento del complejo (CsdA-CsdE) ₂	158
4.4.	Aceptor de azufre del sistema Suf (SufE)	159
4.4.1.	Clonaje de SufE	159
4.4.2.	Expresión de SufE	159
4.4.3.	Purificación de SufE	160
4.4.4.	Experimentos de transferencia de S entre CsdA, CsdE y SufE (<i>crosstalk</i>)	162
4.5.	ARNt treonilcarbamoiladenosina deshidratasa (TcdA)	164
4.5.1.	Clonaje de TcdA	164
4.5.2.	Expresión de TcdA	164
4.5.3.	Purificación de TcdA con cola de histidinas	165
4.5.4.	Clonaje de ARNt ^{Lys} (UUU)	167
4.5.5.	Expresión de ARNt ^{Lys} (UUU)	168
4.5.6.	Purificación de ARNt ^{Lys} (UUU)	169
4.6.	Estudio cristalográfico de TcdA en complejo con ATP y AMP	170
4.6.1.	Cristalización de TcdA en complejo con ATP y AMP	170
4.6.2.	Colección y procesamiento de datos de difracción de rayos-X de los cristales de TcdA en complejo con ATP y AMP	171
4.6.3.	Resolución estructural y refinamiento de TcdA en complejo con ATP y AMP	172
4.7.	Estudio de las interacciones de TcdA con el sistema CSD (CsdA y CsdE)	179
4.7.1.	Caracterización de la unión de TcdA y CsdA por RMN	179
4.7.2.	Caracterización de la unión de TcdA y CsdE por RMN	181
4.7.3.	Caracterización de la unión de TcdA y CsdE por entrecruzamiento	183
4.7.4.	Caracterización de la unión de TcdA y CsdE por SAXS	185
4.8.	Análisis de la interacción de TcdA con ARNt ^{Lys} (UUU)	186
4.8.1.	Ensayos de EMSA (unión de TcdA con ARNt)	186
4.8.2.	Ultracentrifugación analítica del complejo TcdA-tRNA ^{Lys} (UUU)	187
4.8.3.	SAXS del complejo TcdA- ARNt ^{Lys} (UUU)	190
4.9.	PRPP sintetasa (PrsA) de <i>E. coli</i>	192
4.9.1.	Clonaje de PrsA	192
4.9.2.	Expresión de PrsA	192
4.9.3.	Purificación de PrsA	193
4.9.4.	Cristalización de PrsA en complejo con ATP	195
4.9.5.	Colección y procesamiento de datos de difracción de los cristales de PrsA a comple- jados con ATP	195
4.9.6.	Procesamiento de los datos, resolución estructural y refinamiento de PrsA en complejo con ATP	196
4.9.7.	Estructura cristalográfica de <i>EcPrsA</i> en complejo con ATP	197
4.10.	<i>EcOPRT</i> y <i>MtOPRT</i>	201
4.10.1.	Clonaje de <i>EcOPRT</i>	201
4.10.2.	Expresión de <i>EcOPRT</i>	201

4.10.3.	Purificación de <i>EcOPRT</i>	202
4.10.4.	Clonaje de <i>MtOPRT</i>	204
4.10.5.	Expresión de <i>MtOPRT</i>	205
4.10.6.	Purificación de <i>MtOPRT</i>	205
4.10.7.	Cinéticas de <i>EcOPRT</i> y <i>MtOPRT</i>	207
4.10.7.1.	Cinéticas de <i>EcOPRT</i>	207
4.10.7.2.	Cinéticas de <i>MtOPRT</i>	209
4.10.8.	Cristalización de <i>EcOPRT</i> en complejo con 5-FOA y 5-AOA y <i>MtOPRT</i> en su forma apo	210
4.10.9.	Colección de datos y procesamiento de datos de los cristales de <i>EcOPRT</i> en complejo con 5-FOA y 5-AOA y <i>MtOPRT</i> en su forma apo	210
4.10.10.	Resolución estructural y refinamiento de <i>EcOPRT</i> en complejo con 5-FOA y 5-AOA ...	212
4.10.11.	Estructura cristalográfica de <i>EcOPRT</i> en complejo con análogos 5-FOA y 5-AOA	213
4.10.11.1.	Presencia de sulfatos en el centro activo de <i>EcOPRT</i>	216
4.10.12.	Resolución estructural y refinamiento de <i>MtOPRT</i> en su forma apo	217
4.10.13.	Estructura tridimensional de <i>MtOPRT</i> en su forma apo	219
4.10.14.	Superficie de dimerización de <i>MtOPRT</i>	221
5.	Discusión	225
5.1.	Mecanismo de transpersulfuración en el sistema CSD de <i>E. coli</i>	225
5.2.	La interacción de CsdA y CsdE es una unión dinámica	228
5.3.	CsdA es capaz de interaccionar con componentes de los sistemas CSD y Suf	229
5.4.	Interacción transitoria entre CsdE y CsdA/TcdA	231
5.5.	TcdA tiene la tipología de la familia ThiF pero carece de actividad de adenilación y del dominio de unión a zinc	234
5.6.	TcdA presenta un nuevo modo de unión a ARN ^t _{Lys} (UUU), cuyo complejo ha podido ser estudiado mediante diferentes técnicas	236
5.7.	Estructura de la PRPP sintetasa (PrsA) de <i>E. coli</i>	241
5.8.	Estructura de <i>EcOPRT</i> en complejo con 5-FOA y 5-AOA	243
5.9.	Comparación estructural de <i>MtOPRT</i> y <i>EcOPRT</i>	245
5.10.	<i>EcOPRT</i> y <i>MtOPRT</i> son capaces de metabolizar el análogo de OA, 5-FOA	246
6.	Conclusiones	251
7.	Conclusions	255
8.	Referencias	259
9.	Anexo	275
10.	Publicaciones	277

ÍNDICE DE TABLAS

Tabla 1.1.	Estado de oxidación de los diferentes <i>clusters</i> [Fe-S] encontrados en seres vivos	32
Tabla 1.2.	Posiciones en la molécula de ARN ^t portadoras de modificaciones que impliquen átomo de azufre	34
Tabla 1.3.	Sistemas conocidos capaces de captar átomos de azufre para diferentes procesos metabólicos mediante la intervención de la actividad cisteína desulfurasa	36
Tabla 1.4.	Tabla con las cisteína desulfurasas de los sistemas modelo <i>A. vinelandii</i> y <i>E. coli</i>	41
Tabla 3.1.	Cepas de <i>E. coli</i> empleadas en esta tesis	81
Tabla 3.2.	Antibióticos empleados durante el proceso de clonaje y expresión de proteínas recombinantes y ARN ^t	83
Tabla 3.3.	Composición de las soluciones del PCT	105
Tabla 3.4.	Clasificación de los sistemas cristalinos en función de los elementos de simetría	109

Tabla 4.1.	Datos de masas para las especies sin modificar y modificadas (persulfuración) de CsdA	132
Tabla 4.2.	Datos estadísticos de las estructuras de CsdA (estado persulfurado y S-sulfatado)	136
Tabla 4.3.	Datos estadísticos de la solución de las estructuras de los mutantes de CsdA H120A y K250A	141
Tabla 4.4.	Cálculo de la constante de disociación (K_d) para (CsdA-CsdE) ₂ medido por SPR	153
Tabla 4.5.	Datos estadísticos de la solución de las estructuras del complejo (CsdAper-CsdEper) ₂	158
Tabla 4.6.	Valores de masas del ensayo de espectrometría de masas para la persulfuración de CsdE y SufE	163
Tabla 4.7.	Datos estadísticos de la solución de las estructuras de TcdA acomplejada con ATP y AMP	173
Tabla 4.8.	Fragmentos generados del complejo CsdE-TcdA (BMOE) mediante digestión por tripsinización	184
Tabla 4.9.	Datos estadísticos de la solución de la estructura de EcPrsA acomplejada con ATP	197
Tabla 4.10.	Constantes cinéticas de EcOPRT y StOPRT	208
Tabla 4.11.	Constantes cinéticas de MtOPRT	209
Tabla 4.12.	Datos estadísticos de la solución de las estructuras de EcOPRT acomplejada con 5-FOA y 5-AOA	213
Tabla 4.13.	Datos estadísticos de la solución de la estructura de MtOPRT en su forma apo	218
Tabla 5.1.	Comparación de las masas moleculares estimadas a partir de los datos de SAXS	234
Tabla 5.2.	Masas moleculares teóricas y experimentales calculadas mediante SE-UAC para TcdA, tRNA ^{Lys} (UUU) y el complejo formado entre estos dos componentes	237
Tabla 5.3.	Comparación de las masas moleculares estimadas a partir de los datos de SAXS y SE-AUC	238

ÍNDICE DE FIGURAS

Figura 1.1.	Representación de la molécula de PLP y sus reacciones reversibles con aminas primarias para formar iminas	37
Figura 1.2.	Esquema representativo de las 5 familias de enzimas dependientes de PLP según su plegamiento	39
Figura 1.3.	Mecanismo de desulfuración de las cisteína desulfurasas	40
Figura 1.4.	Funcionamiento del sistema Nif	43
Figura 1.5.	Modelo de transferencia de azufre a través del sistema Suf	45
Figura 1.6.	Representación esquemática del operón CSD junto con el gen codificante para TcdA	47
Figura 1.7.	Estructura cristalográfica del dímero de CsdA de <i>E. coli</i>	48
Figura 1.8.	Alineamiento y superposición de las secuencias y estructuras de CsdE y SufE respectivamente	49
Figura 1.9.	Sistemas de captación de azufre en la bacteria <i>E. coli</i> mediante la actuación de cisteína desulfurasas	50
Figura 1.10.	Reacción de deshidratación llevada a cabo por TcdA de <i>E. coli</i> sobre la base t ⁶ A	51
Figura 1.11.	Esquema de la vía de síntesis de tiamina en <i>E. coli</i>	54
Figura 1.12.	Diagrama topológico representativo de la estructura tridimensional de las proteínas E1 o de tipo E1 y ejemplo de estructuras resueltas	55
Figura 1.13.	Secuencias de unión de nucleótidos de las proteínas de tipo E1 de <i>E. coli</i> TcdA, ThiF y MoeB	56
Figura 1.14.	Figura esquemática de la estructura típica de trébol de un ARNt	58
Figura 1.15.	Estructura del t ⁶ A37 en el ASL del ARNt ^{Lys} (UUU)	60
Figura 1.16.	Experimento de coelución para la detección de posibles interactores de CsdE	62
Figura 1.17.	Reacción llevada a cabo por la PRPP sintetasa	63
Figura 1.18.	Similaridad de secuencia de las PRPP sintetasas de <i>E. coli</i> y <i>B. subtilis</i> en el centro activo	65
Figura 1.19.	Estructuras resueltas de PRPP sintetasas de diferentes organismos	67
Figura 1.20.	Reacción enzimática llevada a cabo por la OPRT de forma dependiente de magnesio	68
Figura 1.21.	Ruta de biosíntesis <i>de novo</i> de pirimidinas	69
Figura 1.22.	Estructuras resueltas de OPRT de diferentes organismos	71

Figura 3.1.	Representación esquemática del proceso de mutagénesis dirigida mediante <i>Quik Change Sited-Mutagenesis II</i>	84
Figura 3.2.	Diagrama esquemático del método de mutagénesis dirigida mediante empalme y extensión	86
Figura 3.3.	Molécula del <i>crosslinker</i> BMOE	93
Figura 3.4.	Esquema representativo de los principio de SPR	95
Figura 3.5.	Experimento de NMR de tipo NOESY donde se comparan los espectros generados por una proteína marcada con ^{15}N al incubarla con otra proteína sin marcar	99
Figura 3.6.	Modelo generado por SAXS de tres proteínas a partir de los datos obtenidos experimentalmente	100
Figura 3.7.	Diagrama de fases bidimensional	102
Figura 3.8.	Técnicas empleadas en la obtención de cristales de proteína	104
Figura 3.9.	Sincrotrón ALBA de Barcelona	107
Figura 3.10.	Componentes básicos de una determinación cristalográfica por rayos-X y patrón de difracción	108
Figura 3.11.	Expresiones de los indicadores estadísticos empleados en cristalografía de rayos-X de proteínas.....	110
Figura 3.12.	Expresión derivada del coeficiente de correlación de Pearson	111
Figura 3.13.	Expresión de la transformada de Fourier que representa la función de densidad electrónica	112
Ecuación 3.1.	Ecuación de Patterson	113
Figura 3.14.	Expresión matemática de los <i>R-factors</i> (<i>R-work</i> y <i>R-free</i>)	116
Figura 3.15.	Mapa de Ramachandran mostrando las combinaciones permitidas y prohibidas de Φ (Phi) y Ψ (Psi) en el enlace peptídico	117
Figura 4.1.	Purificación mediante IMAC de CsdA	123
Figura 4.2.	Purificación sustractiva de CsdA tras el corte con TEV	124
Figura 4.3.	Purificación mediante filtración en gel de CsdA	125
Figura 4.4.	Análisis mediante gel de agarosa 1% de las PCR de los dos fragmentos de CsdA H120A mediante el método empalme-extensión	126
Figura 4.5.	Análisis mediante gel de agarosa 1% de la PCR correspondiente a la mutagénesis dirigida mediante el método de <i>QuikChange</i> para la obtención del mutante K250A de CsdA	127
Figura 4.6.	Purificación mediante IMAC de CsdA H120A	128
Figura 4.7.	Purificación mediante filtración en gel de CsdA H120A	129
Figura 4.8.	Purificación mediante IMAC del mutante CsdA K250A	130
Figura 4.9.	Purificación mediante filtración en gel del mutante CsdA K250A	131
Figura 4.10.	Imágenes de difracción de CsdA persulfurada y S-sulfatada	134
Figura 4.11.	Estructura tridimensional del dímero de CsdA persulfurado (CsdAper) ₂	137
Figura 4.12.	Estructura del centro activo de CsdA en diferentes estados de oxidación	138
Figura 4.13.	Imágenes de difracción de los mutantes de CsdA H120A y K250A	140
Figura 4.14.	Estructura cristalográfica del dímero (CsdA) ₂ -H120A	142
Figura 4.15.	Estructura cristalográfica del dímero (CsdA) ₂ -K250A	143
Figura 4.16.	Purificación mediante IMAC de CsdE	145
Figura 4.17.	Purificación mediante filtración en gel de CsdE	146
Figura 4.18.	Purificación mediante IMAC del mutante R21E/N22E de CsdE	148
Figura 4.19.	Purificación mediante filtración en gel del mutante R21E/N22E CsdE	149
Figura 4.20.	Análisis por PAGE-SDS de una muestra del último paso de purificación de cada uno de los mutantes de CsdE	150
Figura 4.21.	Sensorgrama de SPR de la interacción entre CsdE-His ₆ nativa y mutantes con CsdA	152
Figura 4.22.	Representación de la estructura de CsdE por RMN (pdb 1ni7)	154
Figura 4.23.	Datos de dispersión por SAXS, modelado molecular y formas de (CsdA-CsdE) ₂ , CsdA ₂ y CsdE	156
Figura 4.24.	Imagen del cristal y de difracción del complejo (CsdA-CsdE) ₂ persulfurado	157
Figura 4.25.	Purificación mediante IMAC de SufE	160
Figura 4.26.	Purificación mediante filtración en gel de SufE	161

Figura 4.27.	Purificación mediante IMAC de TcdA	166
Figura 4.28.	Purificación mediante filtración en gel de TcdA	167
Figura 4.29.	Prueba de expresión de ARNt ^{Lys} (UUU) analizado mediante gel de agarosa 2%	168
Figura 4.30.	Purificación mediante filtración en gel de ARNt ^{Lys} (UUU)	170
Figura 4.31.	Imagen de difracción de un cristal de TcdA acomplejado con ATP	171
Figura 4.32.	Estructura resuelta por cristalografía de rayos-X de TcdA a 1,77 Å de resolución en complejo con ATP	175
Figura 4.33.	Mapas de densidad de electrones mediante rayos-X de los bolsillos de unión a moléculas de ATP y AMP en el centro activo de TcdA, así como las esferas de coordinación de los iones de K ⁺ y Na ⁺	177
Figura 4.34.	Superficie de dimerización de TcdA	179
Figura 4.35.	Espectros 1D ¹ H- ¹³ C y ¹ H- ¹⁵ N HSQC de TcdA doblemente marcada en su interacción con CsdA	180
Figura 4.36.	Estudios mediante RMN de la interacción transitoria entre TcdA y CsdE	182
Figura 4.37.	Análisis mediante PAGE-SDS del entrecruzamiento mediante BMOE	183
Figura 4.38.	Modelo generado mediante SAXS del complejo TcdA-CsdE formado mediante BMOE	185
Figura 4.39.	Análisis mediante EMSA de la interacción TcdA-ARNt ^{Lys} (UUU)	186
Figura 4.40.	Análisis del complejo TcdA-tRNA ^{Lys} (UUU) mediante AUC	189
Figura 4.41.	Modelo estructural del complejo TcdA-ARNt ^{Lys} (UUU) determinado mediante SAXS	191
Figura 4.42.	Purificación mediante IMAC de la PrsA	193
Figura 4.43.	Purificación mediante filtración en gel de PrsA	194
Figura 4.44.	Imagen de difracción de los cristales de EcPrsA en presencia de ATP	195
Figura 4.45.	Estructura tridimensional de EcPrsA formando complejo con ATP	198
Figura 4.46.	Superposición de los dominios N-terminal y C-terminal de EcPrsA	199
Figura 4.47.	Centro activo de EcPrsA en presencia de ATP	200
Figura 4.48.	Purificación mediante intercambio iónico de EcOPRT	203
Figura 4.49.	Purificación mediante filtración en gel de EcOPRT	204
Figura 4.50.	Purificación mediante IMAC de MtOPRT	206
Figura 4.51.	Purificación mediante filtración en gel de MtOPRT	207
Figura 4.52.	Imagen de difracción de los cristales de EcOPRT	211
Figura 4.53.	Imagen de difracción de los cristales de MtOPRT	212
Figura 4.54.	Esquema de la estructura secundaria de EcOPRT	214
Figura 4.55.	Estructura tridimensional de los complejos de EcOPRT con 5-AOA y 5-FOA	215
Figura 4.56.	Densidad electrónica de las moléculas de 5'FOT (5-FOA) y 5'-AO (5-AOA) en el centro activo de EcOPRT	216
Figura 4.57.	Posición de dos moléculas de sulfato en el centro activo de EcOPRT	217
Figura 4.58.	Esquema de la estructura secundaria de MtOPRT	219
Figura 4.59.	Estructura tridimensional de MtOPRT	220
Figura 4.60.	Detalle de la molécula de sulfato en el S-loop en MtOPRT	221
Figura 5.1.	Representación de la estructura del complejo persulfurado (CsdAper-CsdEper) ₂	227
Figura 5.2.	Alineamiento estructural de las dos cisteína desulfurasas de <i>E. coli</i> , CsdA y SufS	228
Figura 5.3.	Variación estructural en la interfaz en el complejo CsdA-CsdE	231
Figura 5.4.	Esquema de las posibles interacciones a través de la molécula de BMOE entre TcdA y CsdE	233
Figura 5.5.	Comparación de los modelos de proteínas E1 o de tipo E1	236
Figura 5.6.	Modelo de interacción TcdA-ARNt	240
Figura 5.7.	Relación entre el sistema CSD de <i>E. coli</i> y la modificación de ct ⁶ A37 (ARNt ^{ANN})	241
Figura 5.8.	Superposición de la OPRT/5'FOA y OPRT/5'AOA con la enzima apo	244
Figura 5.9.	Superposición de MtOPRT y EcOPRT	246

Abreviaturas y símbolos

^{13}C , carbono 12
 ^{15}N , nitrógeno 15
16S ARN ribosómico
 ^{35}S , azufre 35
5-AOA, 5-aminoorótico (*5-aminoorotic*)
5-FOA, 5-fluoorótico (*5-fluoorotic*)
5-FUMP, 5-fluorouridina monofosfato (*5-fluorouridine monophosphate*)
Å, armstrong
A, adenina
A, adenosina
ADN, ácido desoxirribonucleico
AMP, adenosina monofosfato (*adenosine monophosphate*)
ARN, ácido ribonucleico
ARNt, ácido ribonucleico de transferencia
ASL, *anticodon stem-loop*
ATP, adenosina trifosfato (*adenosine triphosphate*)
AUC, ultracentrífuga analítica (*analytical ultracentrifugation*)
BLAST, *Basic Local Alignment Search Tool*
BMOE, Bis(Maleimido) Etano
BrEt, bromuro de etidio
bs, sitio de unión (*binding site*)
C, citosina
C, cisteína
ct⁶A37, *cyclic threonylcarbamoyladenine* (posición 37 ARNt)
CBB, *Coomassie Brilliant Blue*
CC, coeficiente de correlación
Cd²⁺, cadmio
CIA, *Cytosolic-sulfinase desulfinate*
cpm, cuentas por minuto
CSD, *Cysteine Sulfinase Desulfinate*
CTP, citosina trifosfato (*cytosine triphosphate*)
D, aspártico
Da, dalton
dCMP, desoxicitosina monofosfato (*deoxycytosine triphosphate*)
D_{máx}, dimensión máxima
dNTP, desoxidinucleótidos (*deoxynucleoside triphosphate*)
DO₅₉₀, densidad óptica a 590 nm
Dt, coeficiente de difusión translacional
DTT, *dithiothreitol*
EDTA, ácido etilendiaminotetraacético (*ethylenediaminetetraacetic acid*)
EMSA, *Electrophoretic Mobility Shift Assay*
ESRF, *European Synchrotron Radiation Facility*
F, factor de estructura
Fdx, ferredoxina
Fe, hierro
FT, *flow through*

g, fuerza
 G, guanina
 GDP, guanosina difosfato (*guanosine diphosphate*)
 GMP, guanosina monofosfato (*guanosine monophosphate*)
 H, hélice
 HPLC, *High-Performance Liquid Chromatography*
 HSQC, *Heteronuclear Single-Quantum Correlation spectroscopy*
 Hz, hercios
 I(0), intensidad de dispersión a ángulo cero I(0)
 IMAC, *Immobilized Metal Ion Affinity Chromatography*
 IPTG, Isopropil- β -D-1-tiogalactopiranosido
 Isc, *Iron sulfur cluster*
 K, kelvin
 K, lisina
 Kb kilobase
 K_d, constante de disociación
 KDa, kilodalton
 Krpm, kilo revoluciones por minuto
 L, litro
 LB, Luria-Bertani
 LC-MS, *Liquid Chromatography–Mass Spectrometry*
 M, molar
 M, marcador
 m/z, relación maza/carga
 mA, miliamperios
 MAD, método de Difracción Anómala Múltiple (*Multi-wavelength anomalous dispersion*)
 MALDI-TOF, *Matrix-Assisted Laser Desorption/Ionization, Time Of Flight*
 mAU, miliunidades de absorbancia
 mg, miligramos
 MIR, método de Reemplazamiento Isomorfo Múltiple
 mL, mililitros
 mM, milimolar
 MPD, 2-metil-2,4-pentanodiol
 MR, reemplazamiento molecular (*molecular replacement*)
 MS, espectrometría de masas (*mass spectrometry*)
 MW, masa molecular (*molecular weight*)
 N, asparagina
 N, nucleótido
 NCBI, *National Center for Biotechnology Information*
 NCS, *non-crystallographic symmetry*
 NEB, *New England Biolabs*
 ng, nanogramos
 Nif, *Nitrogen fixation*
 nm, nanometros
 NOESY, *Nuclear Overhauser Effect Spectroscopy*
 NTA, ácido de nitrilotriacético
 OA, ácido orótico (*orotic acid*)
 OMP, orotidina monofosfato (*orotidine 5' monophosphate*)

OPRT, orotato fosforribosil transferasa (*orotate-phosphorybosil transferase*)
 PAGE-SDS, *Polyacrylamide Gel Electrophoresis-Sodium Dodecyl Sulfate*
 pb, pares de bases
 PBS, *phosphate buffered saline*
 PCR, *polymerase chain reaction*
 PCT, *Pre-Crystallization Test*
 PDB, *protein data bank*
 PEG-MME
 PEG, polietilenglicol
 PES, polietersulfona
 PGA-LM, *poly- γ glutmic acid low molecular weight*
 pH, *power of hydrogen*
 Pi, fosfato inorgánico
 pI, punto isoeleétrico
 pKa, constante de disociación ácida
 PLP, *pyridoxal 5'-phosphate*
 PMSF, fluoruro de fenilmetilsulfonilo
 PPi, pirofosfato inorgánico
 ppm, partes por millón
 PRPP, fosforribosil pirofosfato
 PrsA, PRPP sintetasa
 R, arginina
 R5P, ribosa 5-fosfato
 Rg radio de giro
 R_{max}, respuesta máxima teórica
 RMN, resonancia magnética nuclear
 RMSD, *Root-Mean-Square Deviation*
 rpm, revoluciones por minuto
 RU, unidades de respuesta
 S, azufre
 S, coeficiente de sedimentación
 SAD, Difracción Anómala a Longitud Única
 SAXS, *Small Angle X-ray Scattering*
 SDS, *Sodium Dodecyl Sulfate*
 SE, equilibrio de sedimentación
 SEC, *Size Exclusion Chromatography*
 SH⁻, sulfhidrilos
 SOC, *Super Optimal broth with Catabolite repression*
 SPR, *Surface Plasmon Resonance* o Resonancia de plasmón de superficie
 Suf, *Sulfur formation*
 SV, velocidad de sedimentación
 t⁶A37 N6-treonilcarbamoiladenosina
 TCA, ciclo ácidos tricarboxílicos (*tricarboxilic acid cycle*)
 TCD, *tRNA threonylcarbamoiladenosine dehydratase*
 TcdA, *tRNA threonylcarbamoiladenosine dehydratase*
 TCEP, *tris(2-carboxyethyl)phosphine*
 TEV, *Tobacco Etch Virus*
 Tm, temperatura de *melting*

TMP, timidina monofosfato (*thymidine monophosphate*)
U, uracilo
U, uridina
UDP, uracilo difosfato (*uracil diphosphate*)
UMP, uracil monofosfato (*uracil monophosphate*)
UNIPROT, *Universal Protein Resource*
UTP, uracilo trifosfato (*uracil triphosphate*)
UV, ultravioleta
UV-VIS, ultravioleta visible
V, voltios
Ve, volumen elución
W, lavado (*wash*)
YPD, *Yeast extract, Peptone and Dextrose*

Resumen

Introducción

La captura de azufre desde la L-Cys y su movilización son procesos esenciales para la viabilidad celular. Algunos de los componentes biológicos que necesitan de azufre son los *clusters* [Fe-S], los cuales se encuentran involucrados en procesos celulares cruciales. La cisteína desulfurasa es una enzima capaz de extraer los átomos de azufre del sustrato L-Cys y así donarlo a proteínas aceptoras. En la bacteria *Escherichia coli* se encuentran tres sistemas de movilización de azufre dependientes de la actividad cisteína desulfurasa (Isc, Suf y CSD). El sistema Isc es el sistema constitutivo, mientras que Suf es el sistema que actúa bajo condiciones de estrés oxidativo o por falta de hierro. En esta tesis doctoral, el último de los sistemas, CSD, el menos conocido de los tres, ha sido estudiado en profundidad.

El sistema CSD se compone de las proteínas CsdA y CsdE, codificadas por el operón *csdAE*. La cisteína desulfurasa del sistema CSD, CsdA, es una enzima homodimérica dependiente de PLP, la cual actúa como un donador de azufre, mientras que CsdE es la proteína que capta los átomos de azufre de la desulfurasa, estimulando la actividad de ésta última. Estas dos proteínas forman un heterotetrámero estable. La transferencia de azufre a través del sistema CSD involucra un primer paso de desulfuración de la L-cisteína, resultando en la persulfuración de la Cys358 de CsdA, y en un segundo proceso por el cual el persulfuro es cedido a la Cys61 de CsdE vía reacción de transpersulfuración. CsdE persulfurada es capaz de transferir el átomo de azufre a otras dianas.

En adición, hemos estudiado proteínas claves que interaccionan con las proteínas del sistema CSD. La más destacada de estas proteínas es TcdA (anteriormente conocida como CsdL), una treonilcarbamoiladenosina deshidratasa. TcdA está codificada por un gen adyacente al operón CSD y transcrito en dirección opuesta. TcdA cataliza la ciclación de la base t⁶A, modificación encontrada en moléculas de ARN de transferencia (ARNt). La modificación ciclada, ct⁶A, predominantemente encontrada en la base A37 del *stem-loop* del anticodón de ARNts que reconocen codones de tipo ANN, cumple funciones de fidelidad y eficacia en el proceso de traducción.

Otra proteína que ha sido descrita como un interactor del sistema CSD, vía CsdE, es la PRPP sintetasa, PrsA, enzima esencial en la biosíntesis de PRPP, elemento crucial en la síntesis de purinas, pirimidinas, NAD(P) o los aminoácidos triptófano e histidina. PrsA es capaz de metabolizar ribosa 5-fosfato y ATP, con Mg²⁺, en PRPP y fosfato inorgánico. La siguiente enzima en la ruta de síntesis *de novo* de pirimidinas, la orotato fosforribosil transferasa (OPRT), emplea PRPP y ácido orótico (OA) como sustratos. Las OPRTs de *E. coli* y *Mycobacterium tuberculosis* han sido estudiadas como potenciales dianas terapéuticas.

Resultados

Sistema CSD de *E. coli*

- Clonaje, expresión y purificación de CsdA y CsdE (proteínas nativas y mutantes): tanto las proteínas nativas, como los distintos mutantes ensayados en esta tesis fueron clonados, expresados y purificados con éxito.
- Persulfuración de CsdA (autopersulfuración): para la autopersulfuración de la CsdA, ésta fue incubada con L-Cys. Mediante espectrometría de masas fue posible detectar y estudiar esa captación de átomos de azufre por parte de la proteína. Los mutantes expresados de CsdA (H120A y K210A) también se mostraron activos.
- Cristalización de CsdA en distintos estados de oxidación: CsdA pudo ser cristalizada en dos estados de oxidación (persulfurada y S-sulfatada) y resolviendo la estructura a una resolución de 2,05-2,38 Å.
- Cristalización de los mutantes de CsdA (H120A y K250A): los mutantes pudieron ser cristalizados, resolviendo su estructura a resolución de 1,74 y 1,25 Å, respectivamente.
- Ensamblado del complejo (CsdA-CsdE)₂ y persulfuración y cristalización: el complejo (CsdA-CsdE)₂ fue formado a partir de sus componentes. Una vez obtenido, el proceso de persulfuración se llevó a cabo mediante incubación con L-Cys. Posteriormente cristales del complejo persulfurado fueron obtenidos, resolviendo su estructura a una resolución de 2,49 Å.
- Caracterización del complejo (CsdA-CsdE)₂ mediante SPR: con el uso de esta técnica fue posible caracterizar el tipo de interacción entre las proteínas del complejo, así como detectar los residuos de CsdE implicados en el proceso. Para ello, además de usar las proteínas nativas, se realizaron ensayos con los distintos mutantes de CsdE, calculando una K_d de la interacción.
- SAXS del complejo (CsdA-CsdE)₂: el complejo obtenido, así como los componentes por separado, pudieron ser estudiados mediante la técnica SAXS, permitiendo crear un modelo de todos ellos.
- Clonaje, expresión y purificación de SufE: la proteína nativa fue clonada, expresada y purificada con éxito.
- Crosstalk entre los sistemas CSD y Suf: experimentos de transferencia de átomos de S entre componentes de los sistemas CSD y Suf de *E. coli* demostraron el *crosstalk* entre estos sistemas. Gracias a la espectrometría de masas fue posible detectar esta transferencia tanto desde CsdA persulfurada a SufE, como desde el complejo (CsdA-CsdE)₂ persulfurado.

TcdA

- Clonaje, expresión y purificación de TcdA: la proteína nativa fue clonada, expresada y purificada con éxito.
- Clonaje, expresión y purificación de ARNt^{Lys} (UUU): el gen codificante para el ARNt fue obtenido comercialmente. A partir de las construcciones obtenidas fue posible reproducir el gen y clonarlo en un vector de expresión para *E. coli*.
- Caracterización de la interacción CsdA-TcdA y CsdE-TcdA (RMN): gracias a ensayos de RMN fue posible detectar la interacción entre proteínas del sistema CSD y TcdA descritas anteriormente. Marcando la proteína TcdA (¹⁵N y ¹³C), fue posible detectar una interacción débil con la proteína CsdA. Al realizar el experimento con TcdA sin marcar y CsdE doblemente marcada (¹⁵N y ¹³C) fue posible detectar que residuos de CsdE están involucrados en la interacción con TcdA.
- CsdE-TcdA, entrecruzamiento y SAXS: gracias al entrecruzador específico BMOE, fue posible generar un complejo entre las proteína TcdA y CsdE de forma estable, el cual fue estudiado mediante SAXS. Estos experimentos permitieron generar un modelo del complejo en solución a partir de los datos obtenidos del complejo, así como de los modelos cristalográficos de las proteínas por separado.
- Cristalización TcdA acomplejada con ATP o AMP: la proteína recombinante fue cristalizada con las moléculas ATP o AMP, resolviendo las estructuras a una resolución de 1,77 Å para el primer caso y de 1,80 Å para el segundo. Las estructuras resueltas permitieron describir nuevas características de la proteína nunca vistas hasta el momento.
- Interacción TcdA con ARNt^{Lys} (UUU) (EMSA, AUC y SAXS): mediante los ensayos de EMSA, AUC y SAXS fue posible describir la interacción entre estos dos componentes, gracias a la obtención de un complejo de proteína-ácido nucleico de forma específica. Se pudo definir una estequiometría de 1:1 en el complejo obtenido, formado por el dímero de TcdA y dos moléculas de ARNt. El modelo generado es compatible con la actividad enzimática de TcdA y la posición del ARNt el cual es modificado (t⁶A37).

PRPP sintetasa (PrsA)

- Clonaje, expresión y purificación: la proteína fueron clonada, expresada y purificada con éxito.
- Cristalización de PRPP sintetasa en complejo con ATP: la proteína fue cristalizada con la molécula ATP en el centro activo, resolviendo su estructura a una resolución de 1,77 Å. Esto permitió estudiar que residuos están implicados en la interacción con este sustrato.

OPRT (*Escherichia coli* y *Mycobacterium tuberculosis*)

- Clonaje, expresión y purificación de EcOPRT y MtOPRT: las proteínas fueron clonadas, expresadas y purificadas con éxito.
- Estudios cinéticos de EcOPRT y MtOPRT: se midieron parámetros cinéticos para ambas enzimas mediante ensayos con el sustrato OA y con el análogo 5-FOA.
- Cristalización de EcOPRT acomplejada con 5-FOA o 5-AOA: la proteína recombinante fue cristalizada con los análogos de OA, resolviendo la estructura a una resolución de 1,43 Å para el caso del 5-FOA, y de 1,62 Å para el 5-AOA. Las estructuras permitieron describir nuevas características de la proteína nunca vistas hasta el momento.
- Cristalización de MtOPRT en su forma apo: la proteína recombinante fue cristalizada en su forma apo, resolviendo su estructura a una resolución de 1,54 Å. Esta estructura permitió describir esta estructura nunca observada.

Summary

Introduction

Sulfur extraction from L-cysteine and its mobilization are essential processes for cellular viability. Some of the biological components that need sulfur are the [Fe-S] clusters, which are involved in crucial cellular processes. Cysteine desulfurases is the enzyme that can extract sulfur atoms from the L-cysteine substrate and then donate them to other downstream proteins, known as sulfur acceptors. In the bacterium *Escherichia coli* there are three sulfur mobilization systems that depend on cysteine desulfurases for their function (Isc, Suf and CSD). The ISC is a housekeeping system, while the SUF system is activated under oxidative stress or iron starvation. In this PhD thesis, the last system, CSD, which is the least known of the three, has been studied in depth.

The CSD system protein components, CsdA and CsdE, are encoded by the *csdAE* operon. The cysteine desulfurase of the CSD system, CsdA, is a PLP-dependent homodimeric enzyme which acts as the sulfur donor, while CsdE is the cognate sulfur acceptor of CsdA and stimulates CsdA desulfurase activity. These two proteins form a stable heterotetramer. Sulfur transfer through the CSD system involves a first step of L-cysteine desulfuration, which results in a covalent persulfidic modification attached to CsdA Cys358, and a second process whereby the persulfide is transferred to CsdE Cys61 via a transpersulfuration reaction. Persulfurated CsdE, in turn, should be able to transfer the activated sulfur atom to further downstream targets.

Furthermore, we have studied key enzymes that physically interact with the CSD encoded proteins. The most important among them is TcdA (previously known as CsdL), a threonylcarbamoyladenosine dehydratase. TcdA is encoded by a gene located immediately downstream to *csdAE*, which is transcribed from the opposite DNA strand. TcdA catalyze the cyclization of the t⁶A modification found in tRNA ribonucleosides. The resultant cyclic modification, ct⁶A, is predominantly found in A37 of the anti-codon stem-loop (ASL) of tRNAs for ANN codons, where it fulfills functions in translation fidelity and efficiency.

Another protein that has been described as an interactor of the CSD system via CsdE is the PRPP synthetase PrsA, an essential enzyme for the biosynthesis of PRPP, a basic building block for the synthesis of purines, pyrimidines, NAD(P) or the amino acids tryptophan and histidine. PrsA is able to transform ribose 5-phosphate and ATP, with Mg²⁺, into PRPP and inorganic pyrophosphate. The next pathway enzyme in the *de novo* synthesis of pyrimidines, orotate phosphoribosyl transferase (OPRT), uses PRPP

Summary

and orotate (OA) as substrates. OPRT from *E. coli* and *Mycobacterium tuberculosis* have been studied as potential therapeutic targets.

Results and objectives

System CSD from *E. coli*

- Cloning, expression and purification of CsdA and CsdE (native and mutated proteins): the native proteins and the mutants were cloned, expressed and purified successfully.
- Persulfuration of CsdA (autopersulfuration): in order to autopersulfurate CsdA, it was incubated with L-Cys. Using mass spectrometry, it was possible to detect how the protein takes the S atom. The mutants, CsdA (H120A and K210A), showed activity.
- Crystal structures of CsdA in different states: CsdA was crystallized in two different states (persulfurated and S-sulfated) and their crystal structures were solved at 2.05-2.38 Å resolution.
- Crystal structures of CsdA (H120A y K250A): both mutants were crystallized and their crystal structures were solved at 1.74 Å and 1.25 Å, respectively.
- Assembly of the apo and persulfurated (CsdA-CsdE)₂ complexes and crystal structure: (CsdA-CsdE)₂ was assembled from its components. The process of persulfuration was carried out incubating with L-Cys. Later, the crystals of the persulfurated complex were obtained and the crystal structure was determined to 2.49 Å resolution.
- Characterization of CsdA-CsdE interaction by SPR: using this technique it was possible to characterize the interaction between CsdA and CsdE, and the CsdE residues directly involved in the formation of the complex, using the native proteins and the CsdE mutants. K_d of the interactions were calculated.
- SAXS of (CsdA-CsdE)₂: this complex and the components were studied by SAXS. This allowed to restore the overall shape of the complex and to confirm the authenticity of the complex as seen in the crystal structure.
- Cloning, expression and purification of SufE: the native protein was cloned, expressed and purified successfully.
- Crosstalk between CSD and Suf systems: the experiments of sulfur transfer between the CSD and Suf systems from *E. coli* demonstrated the *crosstalk* between these two systems. Mass spectrometry enabled the detection of transferred sulfur atoms from persulfurated CsdA to SufE, and from the persulfurated complex (CsdA-CsdE)₂ to SufE as well.

TcdA

- Cloning, expression and purification: the native protein was cloned, expressed and purified successfully.
- Cloning, expression and purification of tRNA^{Lys}(UUU): the gene coding for the tRNA was purchased. The tRNA gene was subcloned into a transcription vector to produce large amounts of tRNA in *E. coli*.
- Characterization of the CsdA-TcdA and CsdE-TcdA interactions (NMR): it was possible to detect the physical interactions between proteins of the CSD system and TcdA. TcdA was isotopically labeled for the detection of the weak interaction with CsdA. With CsdE, the binding experiment involved unlabeled TcdA and labeled [¹⁵N,¹³C]-CsdE and permitted to identify the individual CsdE residues most affected by the interaction with TcdA.
- CsdE-TcdA, crosslinking and SAXS: using the crosslinker BMOE, capturing a covalently-linked TcdA-CsdE complex and studying it by SAXS was possible. These experiments allowed to restore the shape of the complex in solution and to make an atomic model of the complex by fitting the crystallographic structures of both proteins.
- Crystallization of TcdA with ATP or AMP: TcdA was crystallized with ATP and AMP and the two crystal structures were determined at 1.77 Å and 1.80 Å, respectively. These are the first published crystal structures of TcdA.
- Interaction between TcdA with tRNA^{Lys} (UUU) (EMSA, AUC and SAXS): using these three methods we analyzed the interaction between TcdA and tRNA. A stoichiometry of 1:1 was defined for the complex. Together with the crystal structures of TcdA and tRNA, we constructed hybrid models of the TcdA-tRNA interaction that are compatible with the assumed enzymatic function of TcdA and the position of tRNA which becomes modified (t⁶A37).

PRPP synthetase (PrsA)

- Cloning, expression y purification: the protein was cloned, expressed and purified successfully.
- Crystallization of PrsA with ATP: PrsA was crystallized with ATP in the active site and the crystallographic structure was solved to a resolution of 1.77 Å. Studying the residues involved in the reaction with this substrate was possible.

OPRT (*Mycobacterium tuberculosis* y *Escherichia coli*)

- Cloning, expression and purification: the genes encoding the two individual OPRT enzymes were cloned, and the enzymes were expressed and purified successfully.

Summary

- Kinetic studies of *Ec*OPRT and *Mt*OPRT: The kinetic parameters for both enzymes were measured with the substrate OA and the substrate analogue 5-FOA.
- Crystallographic structure of *Ec*OPRT with 5-FOA or 5-AOA: *Ec*OPRT was crystallized and its crystal structure was solved to 1.43 Å in complex with 5-FOA and to 1.62 Å in complex with 5-AOA. The structures allowed the description of new features and shed light on the binding modes of the two OPRT substrate analogues.
- Crystallographic structure of *Mt*OPRT in the apo form: *Mt*OPRT was crystallized in the apo form and its structure was determined to 1.54 Å. This is the first crystal structure solved of *Mt*OPRT.

Introducción

1. Introducción

1.1. Tráfico de azufre en *E. coli*

La capacidad de captar azufre por parte de los seres vivos y su transferencia a diferentes sistemas del metabolismo celular es esencial para la viabilidad celular a través de la actividad enzimática de múltiples complejos proteicos.

El azufre se encuentra formando parte de diferentes moléculas de la célula, algunos tan básicos como los aminoácidos cisteína y metionina. El reciclaje de éste último es básico ya que es el primer aminoácido de cualquier síntesis proteica. El azufre también forma parte de cofactores como pueden ser el glutatión, empleado en el reciclaje de grupos sulfuro; la tiamina, cofactor esencial en el metabolismo de los carbohidratos; el ácido lipoico, cofactor de deshidrogenasas como la α -cetoglutarato deshidrogenasa; el cofactor de molibdeno involucrado en la acción de determinadas oxidasas; o la coenzima A, implicada en la biosíntesis y oxidación de ácidos grasos. Es de destacar su presencia en la composición de moléculas como los ARN de transferencia (ARNt) vitales en la decodificación del genoma (Sekowska *et al.*, 2000). Finalmente el azufre se puede asociar con el hierro en *clusters* conocidos como *clusters* de hierro-azufre [Fe-S], actuando como cofactores en múltiples reacciones enzimáticas.

1.1.1. Tipo de *clusters* [Fe-S]

Junto con el azufre, el hierro es uno de los metales esenciales en el metabolismo celular. Este metal puede actuar como cofactor en procesos tales como la fijación de nitrógeno, fotosíntesis y respiración, además de formar parte del grupo hemo (Ayala-Castro *et al.*, 2008). Todos estos procesos son esenciales, no solo para los organismos, sino para los ecosistemas en sí.

Otra de las funciones que se le pueden atribuir a estos dos elementos, Fe y S, es su capacidad para formar *clusters* [Fe-S]. Éstos son considerados uno de los primeros cofactores usados por organismos vivos. La presencia de *clusters* [Fe-S] se ha observado en todas las ramas de la vida, participando de forma directa o indirecta en procesos tales como la replicación del material genético del organismo, en su reparación (de vital importancia), en su transcripción y en modificaciones de moléculas como ARN, así como en metabolismo energético (Trotter *et al.*, 2009). Por otro lado, no solo los *clusters* [Fe-S], sino también la movilización del átomo de S en sí, se encuentran directamente relacionados con distintas modificaciones de ARNt (Esberg *et al.*, 1999), como se comentará más adelante.

Introducción

Los *clusters* [Fe-S] están constituidos por combinaciones de diferentes proporciones de átomos de hierro y azufre (Johnson *et al.*, 2005), caracterizados por su capacidad de deslocalizar la densidad electrónica (Sharma *et al.*, 2014), lo que les hace ideales para el transporte de electrones en procesos tales como la respiración, citada anteriormente. Otra de las funciones atribuidas a los *clusters* [Fe-S] es su capacidad de controlar los niveles de Fe y S libres en la célula, los cuales son altamente reactivos, lo que a su vez les puede hacer altamente tóxicos. Por lo tanto, se pueden considerar como método de almacenamiento de estos elementos. En la tabla 1.1 se representan estos tipos de *clusters*.

Estado de oxidación	Ejemplo	Estado de oxidación	Ejemplo
[2Fe-2S] ²⁺	Ferredoxina	[4Fe-4S] ²⁺	Heterodisulfido reductasa
[2Fe-2S] ⁺	Ferroquelatasa	[4Fe-4S] ⁺	Aconitasa
[3Fe-4S] ⁺	Hidrogenasa	[4Fe-4S] ⁰	Sulfito reductasa
[3Fe-4S] ⁰	Hidrogenasa	[8Fe-8S] ⁵⁺	Molibdenum reductasa
[3Fe-4S] ⁻	Hidrogenasa	[8Fe-8S] ⁴⁺	Ferredoxina
[3Fe-4S] ²⁻	Hidrogenasa	[8Fe-7S] ³⁺	Nitrogenasa
[4Fe-4S] ³⁺	Endonucleasa III	[8Fe-7S] ²⁺	Molibdenum nitrogenasa

Tabla 1.1. Estado de oxidación de los diferentes *clusters* [Fe-S] encontrados en seres vivos. Se muestran los diferentes estados de oxidación encontrados en *clusters* [Fe-S] y ejemplos de cada uno (modificación de figura de Johnson *et al.*, 2005).

Los *clusters* [Fe-S] más comunes encontrados en los seres vivos son [2Fe-2S] y [4Fe-4S], normalmente coordinados a través del aminoácido cisteína (Moullis *et al.*, 1996). Generalmente hay un grupo cisteinato (procedente de una cisteína) que completa la coordinación tetrahédrica del hierro pero en ocasiones, se ha observado la sustitución del residuo de cisteína por aminoácidos como aspartato, histidina, serina o incluso el

grupo amida del enlace peptídico, aunque son casos más excepcionales (Johnson *et al.*, 2005).

Además de su función crucial en procesos biológicos como fijación de nitrógeno, fotosíntesis y respiración, se hallan implicados en otras actividades como:

i) El residuo de cisteína que coordina al átomo de Fe del *cluster* [4Fe-4S] facilita la unión de sustratos que sufrirán una deshidratación/hidratación en un amplio espectro de deshidratasa/hidratasa (Flint y Allen, 1996), siendo un ejemplo en este caso la aconitasa. Esta enzima es capaz de convertir el citrato en isocitrato mediante una isomerización estereo-específica.

ii) Otra forma en la actuación de estos *clusters* [Fe-S] es la incorporación de un metal distinto al Fe. Es el caso del *cluster* [Ni-4Fe-5S]. Éste ha sido observado en la monóxido de carbono deshidrogenasa, una oxidorreductasa que emplea CO y H₂O con el fin de reducir diferentes aceptores (Dobbek *et al.*, 2001), y que incluye ese átomo de níquel como parte del cofactor. Un tercer mecanismo es la unión de uno de los Fe del *cluster* [4Fe-4S], mediante un puente cisteínico de un residuo de cisteína. Un ejemplo en este caso sería la Fe-hidrogenasa (Huynh *et al.*, 2014), encargada de la oxidación del H₂. También se han relacionado algunos de estos *clusters* [Fe-S] con funciones estructurales en casos como la endonucleasa III o en proteínas con dedos de zinc (Johnson *et al.*, 2005).

iii) Una función destacable de las enzimas dependientes de este tipo de cofactor son aquellas capaces de modificar moléculas de ARNt. Algunos ejemplos son las enzimas MiaB y Tyw1, ambas encargadas de modificar la posición 37 de este tipo de moléculas, para la tiometilación y para síntesis de la base wybutosina, respectivamente (Alfonzo y Lukes, 2011).

1.1.2. Modificaciones de nucleósidos, vitaminas y cofactores

El átomo de azufre se incorpora mediante la modificación en diferentes biomoléculas como nucleósidos, vitaminas o cofactores.

i) En el caso de los nucleósidos, se ha observado la presencia de azufre de la secuencia de moléculas de ARNt en las posiciones 8, 32, 34 y 37 (Shigi, 2014). En la tabla 1.2 se muestran algunas de las modificaciones descritas en estos puntos que implican átomos de azufre.

Posición ARNt	Modificación	Origen S
8	4-tiouridina (s ⁴ U)	-
32	2-tiocitidina (s ² C)	<i>cluster</i> [Fe-S]
34	5-metilaminometil-2-tiouridina (mnm ⁵ s ² U)	-
34	5-carboximetilaminometil-2-tiouridina (cmnm ⁵ s ² U)	-
37	2-metiltio-N ⁶ -isopenteniladenosina (ms ² i ⁶ A)	<i>cluster</i> [Fe-S]

Tabla 1.2. Posiciones en la molécula de ARNt portadoras de modificaciones que impliquen átomo de azufre. Se muestran algunas de las modificaciones más conocidas así como su posición en la molécula, siendo las más relevantes a nivel de decodificación del genoma las posiciones 34 y 37. Se indica el origen del átomo de S, procedente de *clusters* [Fe-S] u otro todavía sin asignar (-).

De estas modificaciones, las que tienen mayor importancia en el proceso de descifrado del código genético son las involucradas en las posiciones 34 (primera base del anticodón) y 37 (base anexa a la tercera posición del anticodón). En cuanto a la procedencia del átomo de azufre que da estas modificaciones existen dos vías definidas. La primera de ellas es independiente de la formación de *clusters* [Fe-S], donde encontraríamos las modificaciones de la posición 8 y 34, mientras que la segunda vía es dependiente de este tipo de *clusters* [Fe-S], donde encontramos las modificaciones de las posiciones 32 y 37 (Shigi, 2014).

ii) Vitaminas como la tiamina. Esta molécula es el cofactor de enzimas como la piruvato deshidrogenasa, enlazando la glucólisis con el ciclo de los ácidos tricarboxílicos (TCA), la 2-oxoglutamato deshidrogenasa, enzima que forma parte del TCA, la transcetolasa, perteneciente a la ruta de las pentosas fosfato y la acetohidroxiácido sintasa, perteneciente a la ruta de biosíntesis de aminoácidos ramificados (Du *et al.*, 2011).

iii) Otra molécula que posee átomos de azufre en su estructura es la S-adenosil-metionina, una de las principales moléculas donadoras de grupos metilos a todo tipo de moléculas como proteínas, ácidos nucleicos y lípidos entre otras (Rodionov *et al.*, 2004).

iv) Otro cofactor portador de átomos de azufre es la coenzima A implicada tanto en la síntesis como en la oxidación de ácidos grasos (Polyak *et al.*, 2011) y por lo tanto involucrada en el mantenimiento de las membranas celulares entre otras funciones. En el caso del cofactor de molibdeno también encontramos átomos de S, estando este cofactor implicado en la acción enzimática de proteínas tales como la sulfito reductasa (involucrada en la cadena de transportes de electrones) o la xantina oxidorreductasa (implicada en el catabolismo de purinas) (Schwarz *et al.*, 2000). Otro cofactor con átomos de azufre en su estructura es el ácido lipoico, implicado en la actividad enzimática de proteínas involucradas en el catabolismo de los carbohidratos como la α -cetoglutarato deshidrogenasa, perteneciente al TCA, o el complejo piruvato deshidrogenasa (Raddatz y Bisswanger, 1997).

1.2. Cisteína desulfurasas: clasificación

Una vez destacado el papel del azufre en el metabolismo celular y su gran implicación en diferentes procesos es necesario entender cómo las células son capaces de transferir este elemento. El principal mecanismo celular para la captación de los átomos de azufre es la actividad enzimática cisteína desulfurasa (E.C: 2.8.1.7). A través de esta actividad, las enzimas denominadas cisteína desulfurasas son capaces de captar el átomo de azufre de residuos de cisteína libre. Este tipo de enzimas, cuyo estado oligomérico es de homodímero, pertenecen al grupo de las aminotransferasas de clase V (Grishin *et al.*, 1995) y son dependientes del cofactor PLP (*pyridoxal 5'-phosphate*) involucrado en innumerables reacciones enzimáticas. Las cisteína desulfurasas son capaces de captar el átomo de azufre desde su sustrato natural, el aminoácido cisteína (desulfuración) y posteriormente, durante el proceso de transpersulfuración, es donado a otra proteína. El destino último puede ser: i) la formación de *clusters* [Fe-S] (Frazzon *et al.*, 2002). Cuando el destino es el ensamblaje de los *clusters* [Fe-S], el paso siguiente es la transferencia del cofactor maduro a su proteína diana (Ayala-Castro *et al.*, 2008); ii) donar el átomo de azufre a moléculas no proteicas, convirtiendo al residuo de cisteína en una fuente para la formación de diferentes tipos de moléculas como la biotina, la tiamina, el ácido lipoico o los tionucleósidos o algunos ARNt que necesitan de este elemento (Mihara y Esaki, 2002).

Hasta la fecha se han descrito cinco vías por los cuales los seres vivos son capaces de captar átomos de azufre mediante sistemas que implican la acción de cisteína desulfurasa. En Bacteria encontramos cuatro de estos sistemas, el sistema Isc (*Iron sulfur cluster*), el sistema Suf (*Sulfur formation*), el sistema Nif (*Nitrogen fixation*) (Johnson *et al.*, 2005) y CSD (*Cysteine Sulfinate Desulfinate*) (Loiseau *et al.*, 2005). Éste último es el menos conocidos de todos ellos. En eucariotas encontramos el quinto sistema denominado CIA

Introducción

(*Cytosolic-sulfinase desulfinase*), el cual actúa acoplado al sistema de ensamblaje ISC (*Iron-Sulfur Cluster*), presente en mitocondrias (Kispal *et al.*, 1999).

A continuación se muestran cada uno de estos sistemas, reflejándose todos ellos en la tabla 1.3. e indicando la cisteína desulfurasa en cada uno de los casos, así como su función principal.

Sistema	Cisteína desulfurasa	Organismo	Función
Isc	IscS	Bacteria	[Fe-S]
Nif	NifS	Bacteria/Cloroplastos	Nitrogenasa/[Fe-S]
CIA/ISC	NFS1	Eucariotas	[Fe-S]
Suf	SufS	Bacteria/Arquea	[Fe-S]
CSD	CsdA	Bacteria	[Fe-S]/ARNt

Tabla 1.3. Sistemas conocidos capaces de captar átomos de azufre para diferentes procesos metabólicos mediante la intervención de la actividad cisteína desulfurasa. Se muestra cada uno de los sistemas conocidos, así como la cisteína desulfurasa de cada uno de ellos. También se indican los grupos en los que han sido descritos y su función. [Fe-S] en este caso se refiere a la formación de los *clusters* [Fe-S]. En el caso del sistema Nif también se encuentra involucrado de forma especial en la maduración de la nitrogenasa. Por su parte el sistema CSD además está involucrado en la maduración de cierto tipo de ARNt.

1.2.1. Cofactor PLP: un mismo cofactor para una miríada de reacciones

El PLP o piridoxal fosfato (ácido [(4-formil-5-hidroxi-6-metilpiridin-3-il) metil fosfónico], también conocido como vitamina B6, es un cofactor de gran versatilidad en la naturaleza, estando asociado con más de 140 reacciones enzimáticas diferentes (Percudani y Peracchi, 2003). Éste interviene en multitud de procesos tales como transaminación, descarboxilación, racemización y ruptura y formación de enlaces carbono-carbono (Di Salvo *et al.*, 2012). Este tipo de reacciones en las que interviene el cofactor PLP están siempre asociadas a rutas donde intervienen compuestos amino,

principalmente aminoácidos, siendo la excepción la enzima glucógeno fosforilasa (Percudani y Peracchi, 2003). Mediante estas diferentes actividades, las enzimas dependientes de PLP se encuentran involucradas en procesos como el metabolismo de glucosa o lípidos, el de aminoácidos, nucleótidos, así como en la síntesis del grupo hemo y la producción de neurotransmisores (Eliot y Kirsch, 2004; John, 1995), destacando su gran importancia en el metabolismo celular.

La molécula de PLP se caracteriza por tener una estructura química básica formada por el heterociclo piridina con varios sustituyentes, incluyendo un grupo aldehído y un grupo hidroxilo, los cuales le conceden parte de su reactividad (figura 1.1). También es característica su capacidad de formar: i) iminas con grupos amino primarios a través del grupo aldehído y ii) intermediarios carboaniónicos, propiedad que le permite romper enlaces químicos (Di Salvo *et al.*, 2012).

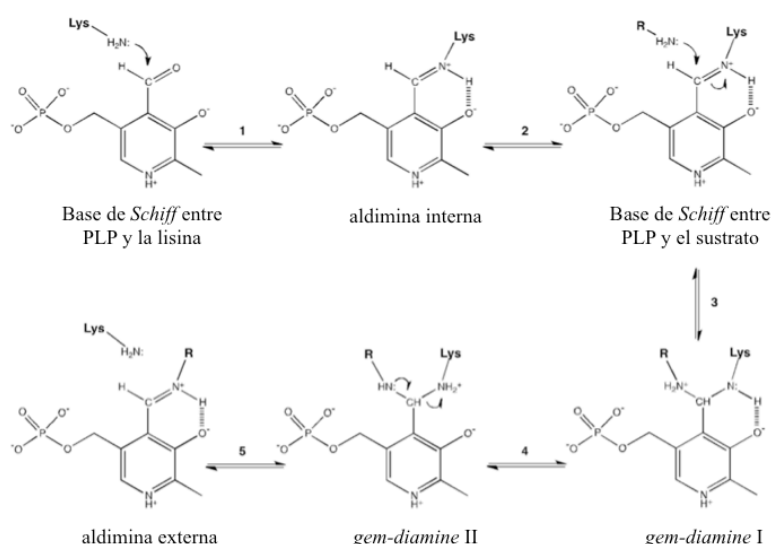


Figura 1.1. Representación de la molécula de PLP y sus reacciones reversibles con aminas primarias para formar iminas. (1) Formación de la aldimina interna con el residuo de lisina del centro activo sobre la molécula de PLP. **(2-5)** Reacción de transaldiminación. La formación de la aldimina externa con el aminoácido sustrato, implica la formación de los intermediarios *geminal diamine* (*gem-diamine*). La reacción de transaldiminación permite la unión y liberación de los sustratos y productos en las reacciones en las que interviene el cofactor (figura modificada de Di Salvo *et al.*, 2012).

Durante la catálisis el cofactor es capaz de formar dos tipos de iminas. La primera imina, es una aldimina interna a través del grupo ϵ -amino de la cadena lateral de un residuo del enzima de lisina totalmente conservado, formando una base de *Schiff*. La segunda imina o aldimina externa, consiste en la unión del cofactor a algún tipo de aminoácido o a otro tipo de compuesto, siendo estas moléculas el sustrato de la reacción. Entre estos dos casos existen formas de la reacción llamados *geminal diamine*. En todas las reacciones llevadas a cabo por enzimas dependientes de PLP la aldimina interna (formada entre la molécula de PLP y el grupo ϵ -amino de la lisina) es sustituida para formar la aldimina externa con el sustrato. Cuando el sustrato es un aminoácido, el grupo amino de éste sustituye al ϵ -amino de la lisina de la proteína (Fernández *et al.*, 2012). Esta aldimina externa es el intermediario común de la reacción a todos los procesos catalíticos llevados a cabo por las enzimas dependientes de PLP.

1.2.2. Familias de enzimas dependientes de PLP

Las proteínas que emplean PLP como cofactor se han dividido en 5 familias (I-V) (figura 1.2), diferenciadas por su distinto tipo de plegamiento (Mehta y Christen, 2000):

- i) En la familia I encontramos las aminotransferasas, las descarboxilasas y las enzimas encargadas de las eliminaciones α , β y γ . En este grupo se encuentran clasificadas las cisteína desulfurasas en estudio.
- ii) La familia II está representada por la enzimas capaces de realizar eliminaciones β .
- iii) En la familia III encontramos la alanina racemasa bacteriana y algunas aminoácido descarboxilasas.
- iv) En la familia IV la D-Ala aminotransferasa y finalmente, en la familia V la glicógeno fosforilasa, así como la almidón fosforilasa (Di Salvo *et al.*, 2012).
- v) En la familia V encontramos la enzima glucógeno fosforilasa.

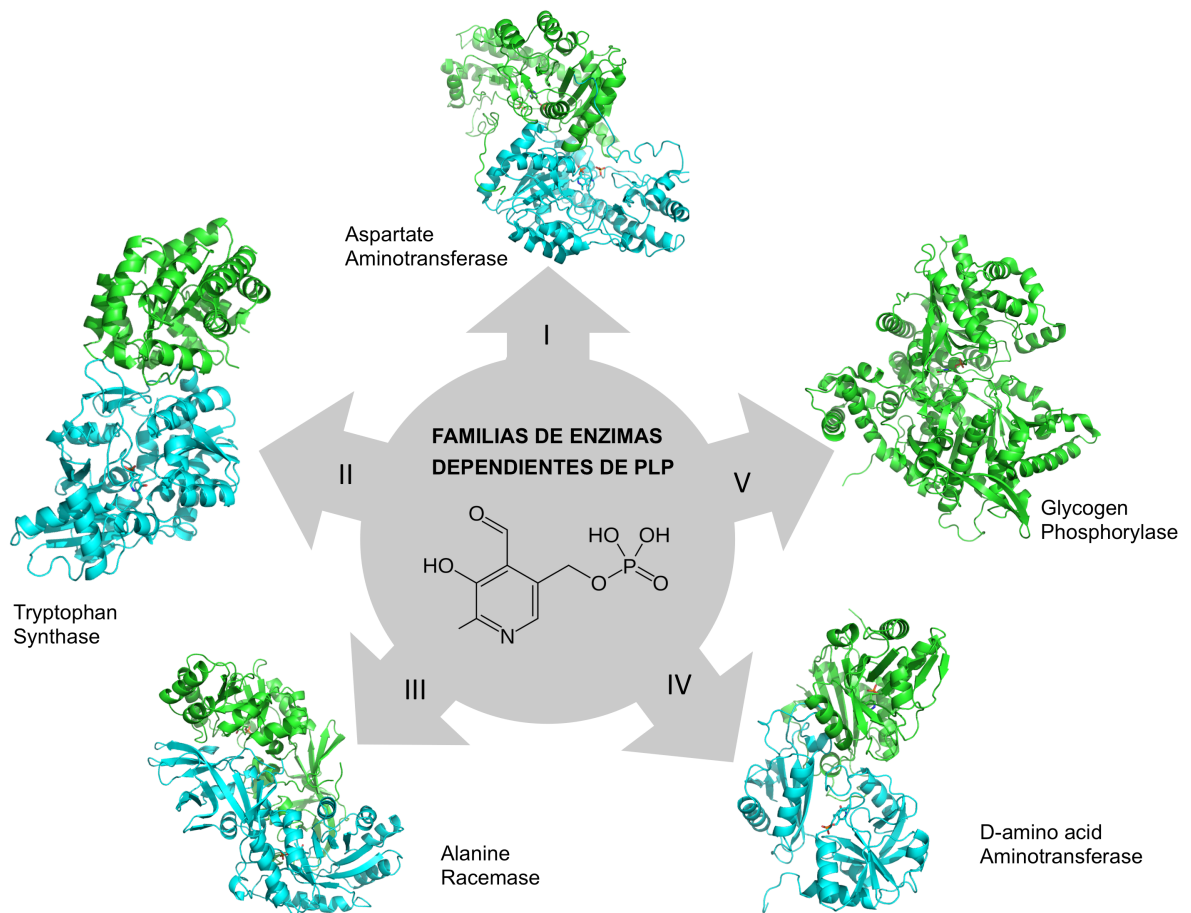


Figura 1.2. Esquema representativo de las 5 familias de enzimas dependientes de PLP según su plegamiento. Se muestran las diferentes familias según el plegamiento, así como la molécula de PLP (zona superior derecha). Dentro de la familia I encontramos entre otras las cisteína desulfurasas.

El hecho de que muchas de las proteínas representadas por las distintas familias citadas arriba, presenten una actividad enzimática similar perteneciendo a diferentes familias según el plegamiento, sugiere un fenómeno de convergencia molecular evolutiva en cuanto a la actividad enzimática debido a necesidades ambientales en un momento dado alrededor de este cofactor (Paiardini *et al.*, 2003).

1.2.3. Mecanismo catalítico de las cisteína desulfurasas

El mecanismo catalítico de las cisteína desulfurasas, está basado en dos pasos de reacción: i) la captura del átomo de azufre (desulfuración) y ii) la cesión de éste a una proteína diana (transpersulfuración) u otro tipo de molécula.

Introducción

En la primera de ellas se da la formación de una base de *Schiff* entre un residuo de cisteína libre (el sustrato) y el cofactor PLP, el cual se encuentra en el centro activo. Aquí también encontramos un residuo de cisteína catalítico de la proteína, la cual se encarga de captar el átomo de azufre desde el sustrato (Ayala-Castro *et al.*, 2008). Así, sobre el grupo tiol de la cisteína catalítica se forma un cisteínil persulfido (Cys-S-SH) (Mihara y Esaki, 2002), entre el átomo de azufre captado (en negrita) y la cisteína catalítica. Esta reacción ocurre gracias al ataque nucleofílico de la cisteína catalítica, liberándose una molécula de alanina como resultado de la desulfuración (figura 1.3). La posibilidad del ataque nucleofílico ocurre también gracias a la actuación del PLP, el cual es capaz de polarizar la molécula de cisteína sustrato, permitiendo así la cesión del grupo tiol para formar el persulfuro (Cys-S-SH).

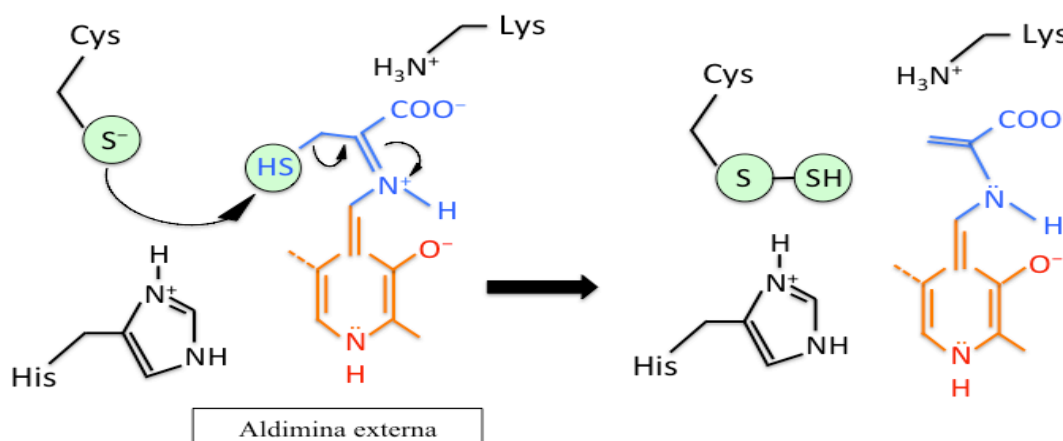


Figura 1.3. Mecanismo de desulfuración de las cisteína desulfurasas. La formación de un grupo persulfido unido a la enzima se da gracias al ataque nucleofílico del tiolato (con carga negativa) del residuo de cisteína del centro activo, sobre el átomo de azufre de la cisteína sustrato (azul) que actúa como donador de este elemento. El ataque nucleofílico se encuentra facilitado gracias a la acción del cofactor PLP (rojo), el cual permite la polarización de la molécula de cisteína sustrato (imagen modificada de Johnson *et al.*, 2005).

Las enzimas cisteína desulfurasas de tipo bacteriano se pueden dividir en dos grupos claramente diferenciados según la secuencia consenso alrededor del residuo de cisteína catalítico que se encarga del proceso de desulfuración y cesión del átomo de azufre (tabla 1.4). Dicha secuencia motivo se halla localizada en el denominado *loop* de

cisteína, en la zona C-terminal de la proteína. Además, como característica diferenciadora, el *loop* cisteína posee una alta movilidad en el grupo I, mientras que en el grupo II, dicho *loop* mantiene una conformación rígida.

En el grupo I, se encuentran NifS e IscS, correspondientes a los sistemas Nif e Isc, respectivamente. Ambos se encuentran presentes en *Azotobacter vinelandii*. En *E. coli*, solo encontramos el segundo de ellos. En el grupo I la secuencia consenso es SSGSACTS, En todos los casos se cita la proteína que mejor representa la familia.

En el grupo II se hallan SufS y CsdA, correspondientes a los sistemas Suf y CSD, respectivamente (en *E. coli*). En este caso la secuencia consenso es RXGXHCA (Ayala-Castro *et al.*, 2008). En la tabla 1.4 se detallan los aminoácidos correspondientes a X en ambos casos).

En concreto, en *E. coli* han sido identificadas tres cisteína desulfurasas, IscS, SufS y CdsA, (Bolstad *et al.*, 2010), pertenecientes a los sistemas Isc, Suf y CSD, respectivamente, que se describirán más adelante.

Cisteína desulfurasa	Organismo	Grupo	Secuencia consenso
NifS	<i>A. vinelandii</i>	I	³²⁰ SSGSACTS ³²⁷
IscS	<i>A. vinelandii</i>	I	³²³ SSGSACTS ³³⁰
IscS	<i>E. coli</i>	I	³²³ SSGSACTS ³³⁰
CsdA	<i>E. coli</i>	II	³⁵³ RAGQHCA ³⁵⁹
SufS	<i>E. coli</i>	II	³⁵⁹ RTGHHCA ³⁶⁵

Tabla 1.4. Tabla con las cisteína desulfurasas de los sistema modelo *A. vinelandii* y *E. coli* . Se señalan las cisteína desulfurasas descritas en los organismos, así como el grupo al que pertenecen según la secuencia consenso alrededor de las cisteína catalítica, letra C negrita. La secuencia consenso de todos los casos se encuentra en la zona C-terminal de las secuencias proteicas (tabla modificada de la mostrada en Mihara y Esaki, 2002).

1.2.4. Reacción de transpersulfuración

El siguiente paso llevado a cabo por este tipo de enzimas es la transpersulfuración, reacción muy similar pues se produce la cesión del átomo de azufre captado por la desulfurasa a una proteína diana sobre la que se ensamblará el *cluster* [Fe-S] o se cederá a otra molécula. En este caso se da de nuevo un ataque nucleofílico sobre el persulfuro por parte de las cisteínas catalíticas que también poseen estas proteínas aceptoras en su centro activo (Mueller *et al.*, 2006). Este sistema de transporte es de alta eficacia, lo que a su vez le vale a la célula para controlar la liberación de grupos de tipo bisulfuro (SH⁻), de alta toxicidad.

1.3. Sistemas enzimáticos bacterianos relacionados con el tráfico de azufre

Como se ha comentado anteriormente en Bacteria existen 4 sistemas involucrados en el tráfico de azufre a partir de un residuo de cisteína, el cual actúa como sustrato inicial de este elemento. A continuación, se describen estos sistemas.

1.3.1. Sistema Isc (*Iron sulfur cluster*)

Este sistema se encuentra conservado tanto en eubacterias como en eucariotas, y en el caso de bacterias es el encargado de actuar bajo condiciones ambientales normales (Singh *et al.*, 2013), siendo el sistema constitutivo. Se ha observado que la inactivación del sistema Isc actúa negativamente en cuanto a la biogénesis de todos los sistemas Fe/S del organismo (Schwartz *et al.*, 2000). Su actividad se centra básicamente en la captación de átomos de azufre para posteriormente ensamblar los citados *clusters* [Fe-S] (Barras *et al.*, 2005). Este sistema se encuentra codificado dentro del operón Isc (Ayala-Castro *et al.*, 2008). La cisteína desulfurasa del sistema Isc (IscS) capta el átomo de azufre de un residuo libre de cisteína para donarlo en la formación de los *clusters* [Fe-S], gracias a las reacciones de desulfuración y transpersulfuración a través de la cisteína catalítica, Cys328. Del mismo modo, en el operón también encontramos los genes codificantes para las proteínas IscU e IscA, las cuales actúan como plataformas para el ensamblaje de los *clusters* [Fe-S], recibiendo los átomos de cisteína procedentes de la IscS. En concreto, los átomos procedentes de la desulfurasa los recibe IscU a través de su residuo Cys63. Estos *clusters* son transportados finalmente a la apoproteína receptora de los cofactores de este tipo. Además, el operón Isc también incluye genes que codifican chaperonas como HscA y HscB, además de Fdx, una proteína con *cluster* [2Fe-2S] de tipo ferredoxina. Por último, tenemos IscX, la cual parece donar los átomos de

hierro para los *clusters* e IscR, el factor de transcripción que regula el operón (Shimomura *et al.*, 2005).

Entre otras funciones, este sistema se encarga de la donación de azufre para la formación de tiouridinas en diferentes posiciones de ARNt (Kessler, 2006), así como para la ruta de síntesis de tiamina y biotina, a través de la enzima ThiI y BioB, respectivamente (Ayala-Castro *et al.*, 2008). Se ha observado que defectos en este sistema ocasionan graves problemas en el crecimiento de *E. coli*, y siendo letal incluso en el caso de *A. vinelandii* (Ayala-Castro *et al.*, 2008).

1.3.2. Sistema Nif (*Nitrogen fixation*)

Este sistema es propio de bacterias capaces de fijar el nitrógeno, como es el caso de *A. vinelandii*, aunque pueden aparecer homólogos en organismos totalmente diferentes, habiéndose observado en casos tales como *Helicobacter pylori* (Tokumoto *et al.*, 2004) y plantas. En el caso de plantas se encuentran en los cloroplastos, donde su deficiencia altera notablemente el proceso de fotosíntesis (Hoewyk *et al.*, 2007), en concreto la producción de ferredoxina y el fotosistema I (Yabe *et al.*, 2004), este último es imprescindible en el proceso de transporte de electrones de la fotosíntesis.

En el sistema Nif, la proteína NifS es la cisteína desulfurasa, capaz de captar los átomos de azufre desde una molécula de cisteína sustrato empleando su residuo catalítico, la Cys325, mediante la reacción de desulfuración. Posteriormente, gracias a la transpersulfuración, se cede el átomo de azufre a NifU. Esta proteína además también es capaz de unir un átomo de hierro, actuando como una plataforma para la formación de los *clusters* [Fe-S] (figura 1.4). A continuación, estos *clusters* son cedidos a la nitrogenasa para que pueda llevar a cabo su función fijadora de nitrógeno en el caso de *A. vinelandii* (Johnson *et al.*, 2005) o a otras proteínas diana que intervengan en procesos como la fotosíntesis (Yabe *et al.*, 2004).

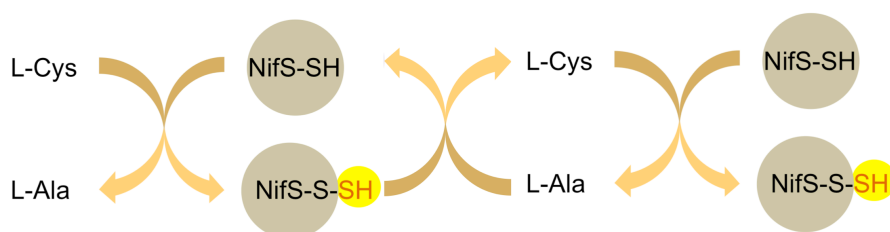


Figura 1.4. Funcionamiento del sistema Nif. Se muestran los diferentes componentes del sistema Nif (NifS y NifU), mostrándose en rojo el grupo tiol que es transportado desde la cisteína sustrato hasta los *clusters* [Fe-S].

1.3.3. Sistema Suf (*Sulfur formation*)

El sistema Suf actúa en situaciones de falta de hierro o de estrés por oxidación (Ayala-Castro *et al.*, 2008). En este caso encontramos este sistema tanto en bacterias como en arqueas (Iwasaki, 2010), además de haber sido observado en el parásito de la malaria, causante del paludismo (Bolstad *et al.*, 2010). El sistema Suf se encuentra codificado por el operón Suf, el cual codifica para seis proteínas diferentes. Como en el sistema Isc, encontramos una cisteína desulfurasa, SufS, la cual tras captar el átomo de azufre igual que en los casos anteriores, mediante la reacción de desulfuración a través de la Cys364 (en el caso de *E. coli*), lo dona a la proteína plataforma SufE, mediante la reacción de transpersulfuración, actuando en el caso del aceptor la Cys51 (Ayala-Castro *et al.*, 2008). Estas dos proteínas, SufS y SufE, interaccionan entre sí formando el complejo (SufS-SufE), y cuya formación acelera los procesos de desulfuración y transpersulfuración (Loiseau *et al.*, 2003). A continuación, el complejo SufBC₂D, formado por proteínas también codificadas por el operón (SufB, SufC y SufD), se encarga de captar el átomo de azufre procedente de SufE. Este complejo se ha caracterizado por su capacidad de actuar como plataforma en el ensamblaje de *clusters* [4Fe-4S] (Singh *et al.*, 2013). Una vez formados los *clusters*, la proteína SufA se encarga de transportarlos hasta las apoproteínas dianas para cumplir su función (Chehal *et al.*, 2009). En la figura 1.5 se muestra un esquema de las enzimas implicadas en este sistema.

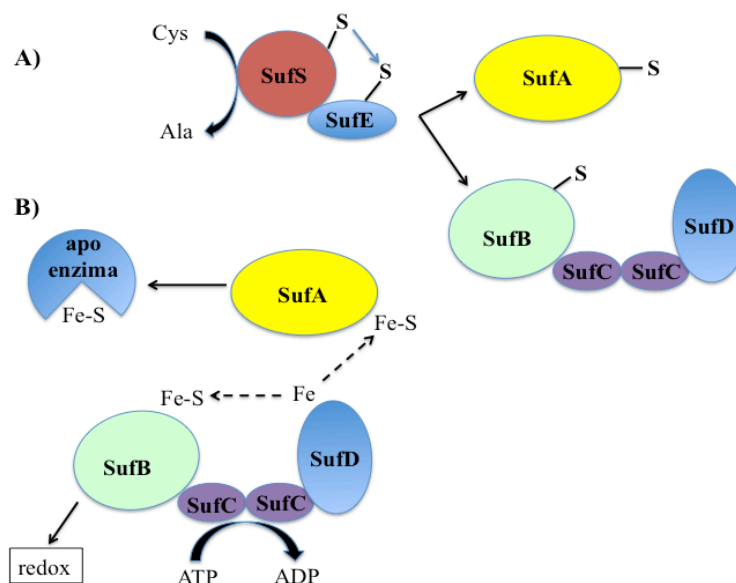


Figura 1.5. Modelo de transferencia de azufre en el sistema Suf. a) SufS y SufE captan y movilizan el átomo de azufre (cisteína libre) para donarlo al complejo SufBC₂D y a SufA. b) Los *clusters* [Fe-S] se pueden formar tanto sobre SufA como sobre SufB, siendo transportados o a apoproteínas o participando directamente en procesos redox. Las líneas discontinuas representan posibles donaciones de hierro mediante SufD o directamente a SufA. El fin de la actividad ATPasa de SufC es desconocido (modificación de Ayala-Castro *et al.*, 2008).

Se ha observado que mutaciones en el gen codificante para SufS no afectan de forma significativa al fenotipo bajo condiciones normales en bacteria, al contrario de lo observado bajo situaciones de estrés, donde la formación de proteínas portadoras de estos *clusters* se ve afectada significativamente (Outten *et al.*, 2004).

Además, siguiendo con la fisiología de bacterias, cabe destacar la importancia de estos dos sistemas (Isc y Suf) en la viabilidad celular, ya que un doble mutante *iscS/sufS* es letal para *E. coli* (Outten *et al.*, 2004). Esto implica que estos dos sistemas de desulfurasas son esenciales en el crecimiento y mantenimiento óptimo celular, aunque también depende del organismo de estudio. Por ejemplo, en el caso de la bacteria modelo *E. coli* el sistema Suf es el de mayor importancia entre los diferentes sistemas a la hora de ensamblar los *clusters* [Fe-S]. Sin embargo, en casos como *Mycobacterium tuberculosis* o ciertas arqueas solo parece existir el sistema Suf, siendo éste posiblemente el constitutivo en estos casos. Por otro lado, en el proceso de la formación de los *clusters* [Fe-S] en bacterias de tipo cianobacteria, el sistema con mayor relevancia es Isc. Para

organismos eucariotas es algo más complejo, debido a que los sistemas de sulfuración se hallan compartimentalizados y en cada orgánulo presenta más prevalencia alguno de los sistemas. Así homólogos del sistema Isc se encuentran representados mayoritariamente en las mitocondrias, mientras que homólogos del sistema Nif se encuentran representados con mayor frecuencia en los cloroplastos (Ayala-Castro *et al.*, 2008).

1.3.4. Sistema CSD (*Cysteine Sulfinase Desulfinase*)

El último sistema bacteriano es CSD, cuya cisteína desulfurasa CsdA se caracteriza por ser, de las tres descritas en *E. coli* (IscS, SufS y CsdA), la que mayor actividad muestra en condiciones *in vitro* (Trotter *et al.*, 2009). CsdA es una cisteína desulfurasa del grupo II (tabla 1.4) que presenta la secuencia motivo ³⁵³RAGQHCA³⁵⁹, donde la cisteína catalítica corresponde al residuo Cys358, encargada de las reacciones de desulfuración y transpersulfuración. En este sistema la proteína CsdE actúa de forma muy similar al caso de SufE, formando un heterotetrámero con la proteína CsdA (CsdA-CsdE)₂ (Loiseau *et al.*, 2005), lo que aumenta en dos veces la capacidad que tiene CsdA por captar y transferir átomos de azufre a CsdE. Además, este sistema es capaz de interactuar, al menos en el caso de *E. coli*, con otros sistemas como Suf, en un proceso denominado *crosstalk*. El sistema CSD ha sido objeto de estudio de la presente tesis como sistema modelo del tráfico de azufre bacteriano.

1.4. Miembros del sistema CSD

Como se ha comentado anteriormente, el sistema CSD se encuentra controlado y codificado por el operón CSD, encontrando dentro del mismo los genes codificantes para la cisteína desulfurasa CsdA y la proteína aceptora de los átomos de azufre, CsdE. En la figura 1.6 se muestran los componentes del operón CSD. Recientemente fue descrita una proteína capaz de interactuar directamente con una de las proteínas de este sistema, la proteína TcdA (conocida anteriormente como YgdL y CsdL) cuyo gen se encuentra anexo al operón en cuestión, aunque transcrito en sentido opuesto, como se puede observar en la figura 1.6. Una vez el complejo donador-aceptor de azufre (CsdA-CsdE)₂ ha capturado el átomo de azufre, lo transfiere a TcdA, la cual se ha descrito como capaz de modificar ciertos ARNts mediante una reacción de deshidratación de la base 37 (Miyauchi *et al.*, 2013).

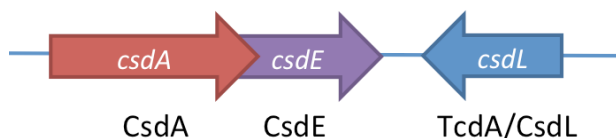


Figura 1.6. Representación esquemática del operón CSD junto con el gen codificante para TcdA. Este operón contiene los genes codificantes para las proteínas CsdA y CsdE. Adyacente a él se encuentra el gen codificante para TcdA, transcrito en la dirección opuesta a los otros dos genes, proteína encargada de modificar cierto tipo de ARNt.

En ensayos *in vivo* del operón CSD, se ha observado que cuando éste es defectuoso no se da ningún efecto aparente en el metabolismo celular en cuanto a enzimas con cofactores formados por *clusters* [Fe-S] se refiere (Trotter *et al.*, 2009). Sin embargo, *in vitro* se ha observado que la pareja CsdA/CsdE participa en la reconstitución de enzimas como [4Fe-4S] NadA y la 6-fosfogluconato deshidratasa (Ayala-Castro *et al.*, 2008), estando la primera de ellas implicada en la síntesis de NAD (Nicotinamida Adenina Dinucleótido) y la segunda formando parte de la ruta de Entner-Doudoroff, ruta energética alternativa a la glucólisis, capaz de metabolizar también la glucosa hasta piruvato.

1.4.1. CsdA, cisteína desulfurasa

La cisteína desulfurasa del sistema CSD es el enzima homodimérica CsdA, dependiente de PLP, perteneciente al grupo II de cisteína desulfurasas según la secuencia consenso (tabla 1.4), alrededor del residuo de cisteína catalítico Cys358. La mutación de Cys358 por serina, elimina totalmente la actividad desulfurasa (Mihara *et al.*, 2000).

Por otro lado, uno de los aceptores de los átomos de azufre de la proteína es CsdE, la cual es capaz de aumentar por dos la actividad desulfurasa de CsdA al formar un complejo con CsdE. El complejo que se forma es un heterotetrámero (CsdA-CsdE)₂, donde el dímero de CsdA forma el núcleo del complejo, acomplándose a cada lado de éste un monómero de CsdE (Loiseau *et al.*, 2005).

La estructura de CsdA de *E. coli* fue resuelta mediante difracción de rayos-X en su forma apo (Kim y Park, 2013) demostrándose su gran similitud estructural con las otras cisteína desulfurasas de esta misma enterobacteria, además de NifS, aunque no es posible visualizar la cisteína catalítica (Cys358). La estructura de CsdA se puede dividir en dos dominios, observándose la región grande (*large region*) y la región pequeña (*small*

region) (Kim y Park, 2013) (figura 1.7). La primera de ellas está formada por 7 hebras β paralelas, flanqueadas por 7 hélices α . La region pequeña consiste en 4 hebras β antiparalelas flanqueadas por 3 hélices α . La molécula de PLP se encuentra en la interfaz de ambas regiones, siendo su lisina de unión la Lys222, responsable de la formación de la base de *Schiff*. En cuanto a la zona de dimerización, los elementos componentes de ésta son bastante similares en los casos de las demás cisteína desulfurasas de *E. coli* (IscS y SufS), denotando un origen común, como se puede deducir al poseer un plegamiento general similar.

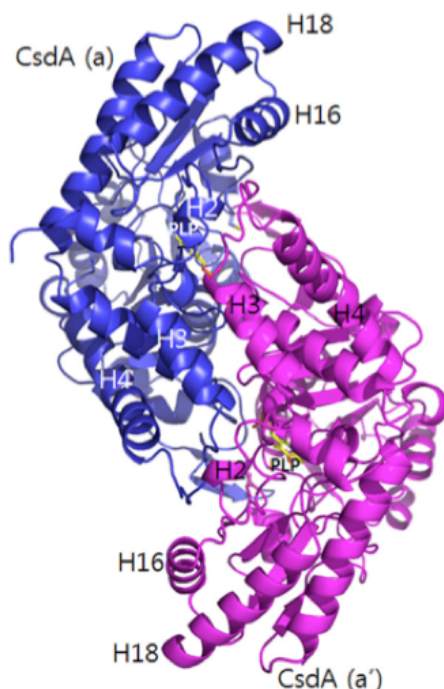


Figura 1.7. Estructura cristalográfica del dímero de CsdA de *E. coli*. Se puede observar cada uno de los monómeros (a y a') con el plegamiento de tipo I citado, así como las dos moléculas del cofactor PLP. La letra H se refiere a la designación de las hélices α de la molécula de CsdA (Kim y Park, 2013).

1.4.2. CsdE, aceptora de átomos de azufre de CsdA

La proteína CsdE, cuyo gen codificante también pertenece al operón CSD, es la encargada de recibir los átomos de azufre desde la proteína CsdA en la reacción, conocida como transpersulfuración, a través de su único residuo de cisteína, la Cys61, para posteriormente donarlo a otros aceptores (Ayala-Castro *et al.*, 2008).

Esta proteína, cuyo estado oligomérico es de monómero cuando no se encuentra formando parte del tetrámero con la desulfurasa (CsdA-CsdE)₂, guarda gran similitud tanto en secuencia como estructuralmente con otros componentes de los otros sistemas mencionados, en concreto con la proteína SufE. Ésta proteína, como en el caso de CsdE, es la encargada de tomar los átomos de azufre de SufS en el proceso de transpersulfuración, en este caso a través de la Cys51 (Bolstad *et al.*, 2010). CsdE y SufE

son proteínas homólogas, guardando un porcentaje de identidad de secuencia entre ellas del 35%, además de mostrar un gran parecido estructural (figura 1.8).

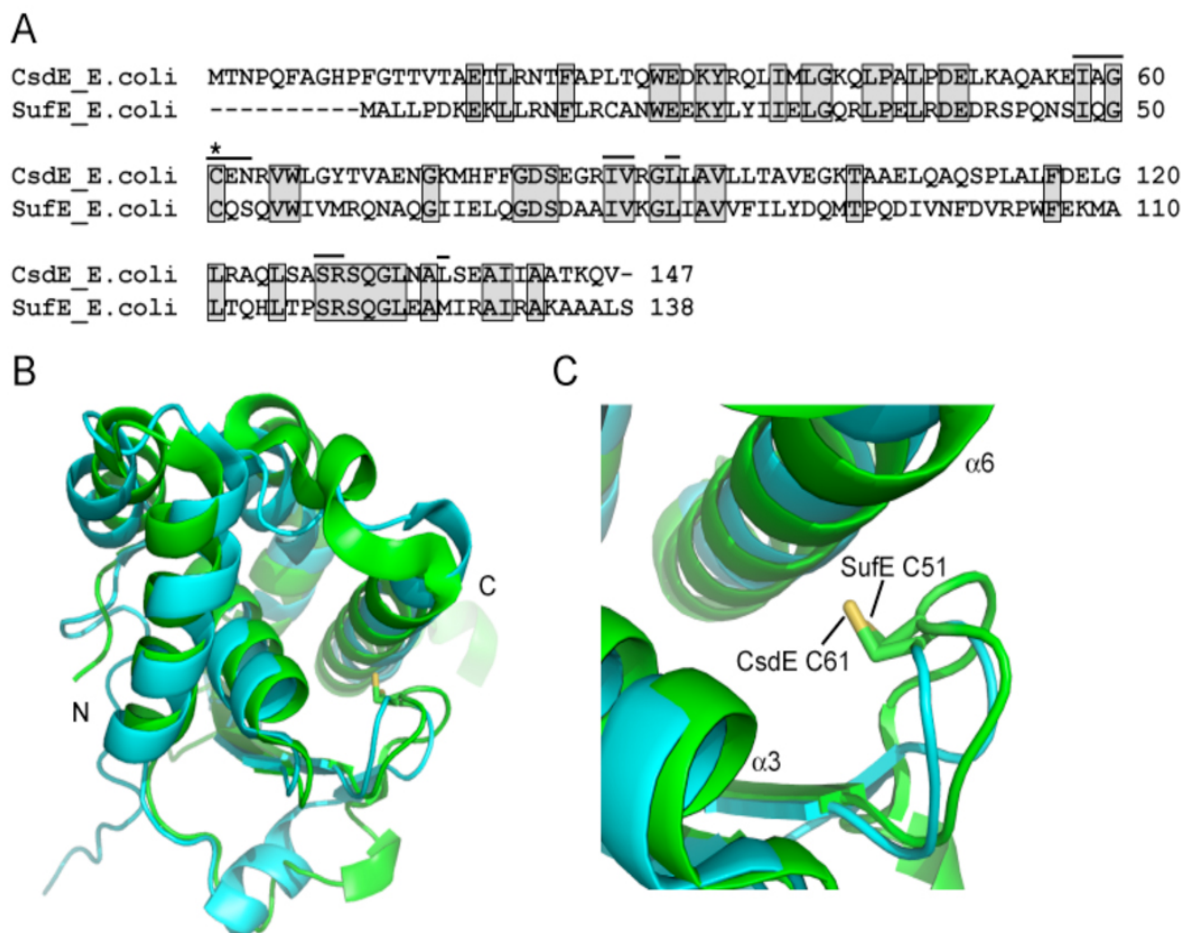


Figura 1.8. Alineamiento y superposición de las secuencias y estructuras de CsdE y SufE respectivamente. **a)** Alineamiento de las secuencias proteicas de ambas proteínas. Los residuos compartidos se muestran sombreados, mientras que las barras indican los residuos que forman parte de la cavidad de la cisteína reactiva, indicada por un asterisco. **b)** Alineamiento de las estructuras de CsdE (pdb 1ni7) y SufE (pdb 1mzg), representándose la primera en cian y la segunda en verde, donde se observa el residuo de la cisteína catalítica, mostrándose los átomos de carbón en verde y los de azufre en amarillo (Bolstad *et al.*, 2010). CsdE fue resuelta por en Liu *et al.*, 2005 y SufE en Goldsmith-Fischman *et al.*, 2004.

La estructura tridimensional de CsdE fue resuelta inicialmente mediante RMN (Resonancia Magnética Nuclear), determinando una estructura de plegamiento α/β , estando compuesta por seis hélices α y 3 hebras β antiparalelas. En la figura 1.8. (b) se puede observar la estructura de la proteína (Liu *et al.*, 2005) y su superposición con

Introducción

SufE. De manera similar a SufS, la cisteína Cys61 ocupa una posición prominente en el monomero, situada en un *loop* relativamente flexible.

Cabe destacar que en el trabajo de Loiseau *et al.*, 2005 se indica que SufE, proteína aceptora del sistema Suf y de gran parecido estructural con CsdE, no es capaz de tomar átomos de azufre de la desulfurasa CsdA (Loiseau *et al.*, 2005), al contrario de otros autores (Trotter *et al.*, 2009) que sugieren que CsdA es capaz de interaccionar con SufE.

Por lo tanto en *E. coli*, encontramos tres sistemas capaces de captar azufre a través de cisteína desulfurasas, siendo uno de ellos el descrito como el constitutivo (Isc), otro el que actúa bajo condiciones de estrés (Suf) y por último el menos conocido de los tres y del cual no se conoce exactamente su función en la fisiología de esta bacteria (CSD) (figura 1.9).

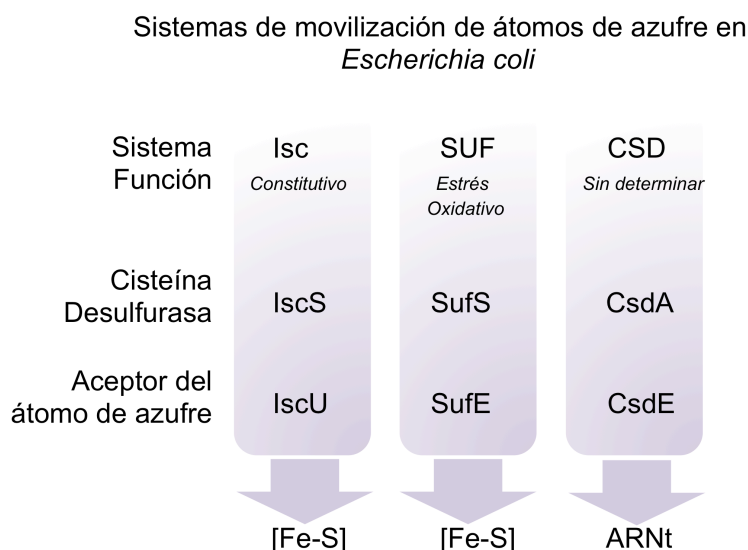


Figura 1.9. Sistemas de captación de azufre en la bacteria *E. coli* mediante la actuación de cisteína desulfurasas. Esquema de los tres sistemas conocidos de esta enterobacteria modelo en la captación de azufre, indicándose las cisteína desulfurasas de cada uno de los ellos, así como las proteínas capaces de recibir ese elemento en la reacción de transpersulfuración. Los dos primeros sistemas están relacionados con la formación de *clusters* [Fe-S] de forma constitutiva (sistema Isc) y en condiciones de estrés (sistema Suf). Por su parte, el sistema CSD puede interactuar con el sistema Suf vía CsdA-SufE (bajo condiciones desconocidas) en el proceso de *crosstalk* o seguir su ruta a través de CsdE para la modificación de ARNt.

1.4.3. TcdA, proteína modificadora de ciertos tipos de ARNt

Esta proteína conocida inicialmente como YgdL y CsdL, por creerse propia del sistema CSD (Trotter *et al.*, 2009), fue llamada posteriormente TcdA, por su actividad enzimática recientemente determinada, además no estar codificada dentro del operón CSD (figura 1.6). En concreto es una deshidratasa, tRNA *threonylcarbamoyladenine dehydratase* (EC: 6.1.-.-). Cataliza la deshidratación dependiente de ATP de la base treonilcarbamoyladenina (t^6A), localizada en la posición 37 de los ARNt capaces de reconocer codones que comienza con adenosina (ANN) (Miyauchi *et al.*, 2013). La molécula resultante de este proceso de deshidratación es la ciclación de la base sustrato, dando como resultado ct^6A , la forma ciclada de t^6A (figura 1.10). En relación al sistema CSD, TcdA es capaz de recibir átomos de azufre tanto desde CsdA directamente, como a través de CsdE (Trotter *et al.*, 2009), relacionando este proceso con el *fitness* de *E. coli*.

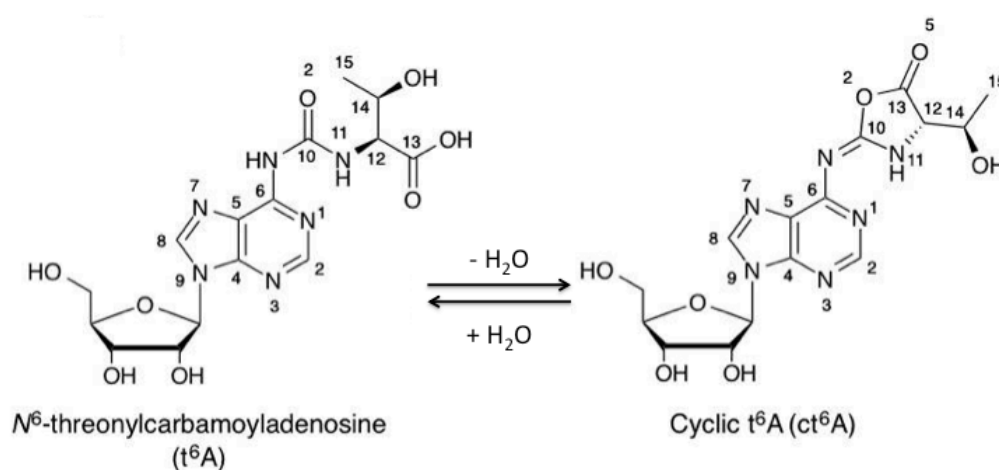


Figura 1.10. Reacción de deshidratación llevada a cabo por TcdA de *E. coli* sobre la base t^6A . La reacción de deshidratación llevada a cabo por la enzima es dependiente de ATP, mientras que la reacción de hidratación ocurre espontáneamente sobre ciertas condiciones ambientales (López-Esteva *et al.*, 2015).

Siguiendo con las funciones de la proteína TcdA, Miyauchi *et al.*, 2013 también demostraron que la ausencia de los genes codificantes para CsdA y/o CsdE puede afectar, de manera significativa, a la ciclación llevada a cabo por TcdA sobre la base t^6A en la posición 37 de distintos ARNts. Parece que la actividad de la deshidratasa depende de la actuación de las anteriores proteínas de una forma directa. Este hecho fue

demostrado cuando se observó una disminución notable de la ciclación de las moléculas diana al ensayar mutantes $\Delta csdE$ y $\Delta csdA$ (Miyachi *et al.*, 2013) de *E. coli*. Sin embargo, en el mismo trabajo se demostró que TcdA es capaz de llevar a cabo su actividad *in vitro*, simplemente con la presencia del sustrato y de ATP (Miyachi *et al.*, 2013).

En contraste, en *E. coli*, la delección para el gen de TcdA no afecta de manera detectable al crecimiento celular, aunque se ha observado que las células no son capaces de crecer bajo condiciones de competitividad con una cepa salvaje. Además, se ha demostrado que la forma cíclica es más efectiva que la no cíclica a la hora de la decodificación por parte ciertos tipos de ARNt, como es el caso de ARNt^{Lys}, además de observarse errores en la lectura de codones *stop* cuando esta ciclación no se da (Miyachi *et al.*, 2013).

1.4.3.1 Homólogos de TcdA en eucariotas

En eucariotas, en concreto en la levadura modelo *Saccharomyces cerevisiae* se han descrito dos homólogos de TcdA, TCD1 y TCD2 (*tRNA threonylcarbamoyladenine dehydratase 1 y 2*, respectivamente). Del mismo modo que en *E. coli*, es decir, mediante el estudio de delecciones de los genes correspondientes y análisis de la proporción de ciclación de la base diana, se observó que la ausencia de esta actividad en la levadura conduce a defectos en el proceso de respiración, sugiriendo que esta modificación es de gran importancia en la decodificación llevada a cabo en las mitocondrias de esta clase de organismos, lugar donde debe de actuar (Miyachi *et al.*, 2013). Adicionalmente también se observó que delecciones para los genes codificantes para estas proteínas generan cepas incapaces de crecer en presencia del medio YPD (*Yeast extract, Peptone and Dextrose*), altamente rico, indicando que alguno de estos componentes implicados (ct⁶A o TCD1 o TCD2) actúan directa o indirectamente en el proceso de crecimiento celular.

Al buscar proteínas similares humanas, encontramos la proteína UBA5 también perteneciente a la superfamilia de tipo E1. UBA5 se encarga de activar de forma dependiente de ATP y zinc, a las proteínas UFM1 y SUMO2 (Schulman y Harper, 2009). A través de esta vía, UBA5 influye en procesos tales como modificación de receptores nucleares o de estrés del retículo endoplasmático por parte de UFM1 (Yoo *et al.*, 2014) o en transporte nuclear, regulación de la transcripción, apoptosis o estabilidad proteica en el caso de SUMO2 (Maarifi *et al.*, 2015).

1.4.3.2. Homólogos de TcdA en procariotas

El análisis de secuencia de la proteína TcdA mediante BLAST (*Basic Local Alignment Search Tool*), identifica a esta enzima dentro de la superfamilia de proteínas de tipo E1, y muestra que TcdA y las proteínas de la superfamilia de proteínas de tipo E1, ThiF y MoeB, son homólogas. En concreto, muestra similitud la zona N-terminal pertenece a esta superfamilia de proteínas, aunque no comparte el dominio C-terminal de unión a zinc. Tanto ThiF como MoeB, son capaces de unir zinc además de ATP para llevar a cabo sus funciones (Duda *et al.*, 2005; Lake *et al.*, 2001).

Al comparar TcdA con sus homólogos en *E. coli*, ThiF y MoeB, se observa que el porcentaje de identidad entre estas proteínas no es particularmente alto, siendo de 29,0% entre el par TcdA/ThiF y del 32,8% en el par TcdA/MoeB. Este porcentaje no es tan alto debido a que en la mitad C-terminal de TcdA no se encuentra el dominio de unión al metal que unen las otras dos, el zinc, aunque tampoco existe un gran porcentaje de identidad en el par MoeB/ThiF, siendo este del 43,5%. La función de ThiF y MoeB es la adenilación dependiente de una molécula de ATP, el extremo C-terminal de las proteínas de tipo ubiquitina (Burrough *et al.*, 2009), ThiS y MoaD, respectivamente. Este suele ser un rasgo típico de las enzimas de la superfamilia de proteínas de tipo E1, además de mostrar también una capacidad de transporte de átomos de azufre.

i) Proteína ThiF

La proteína ThiF se encuentra involucrada directamente en la síntesis de tiamina o vitamina B1. ThiF persulfura y adenila la proteína ThiS. La cisteína catalítica de ThiF es la Cys184, (Duda *et al.*, 2005) residuo no conservado en TcdA. De esta forma, en la proteína ThiF se genera un acildisulfido, lo que desemboca finalmente de la formación de un grupo tiazol (Lehmann *et al.*, 2006), participando en la formación de tiamina (Duda *et al.*, 2005) (figura 1.11). Además, se ha descrito que ThiF es capaz de unir el metal zinc, mediante la coordinación de 4 residuos de cisteína (169, 172, 240 y 243) (Lehman *et al.*, 2006), todas ellas conservados también en MoeB. Este metal es esencial a la hora de llevar a cabo la actividad enzimática de ambas proteínas.

Introducción

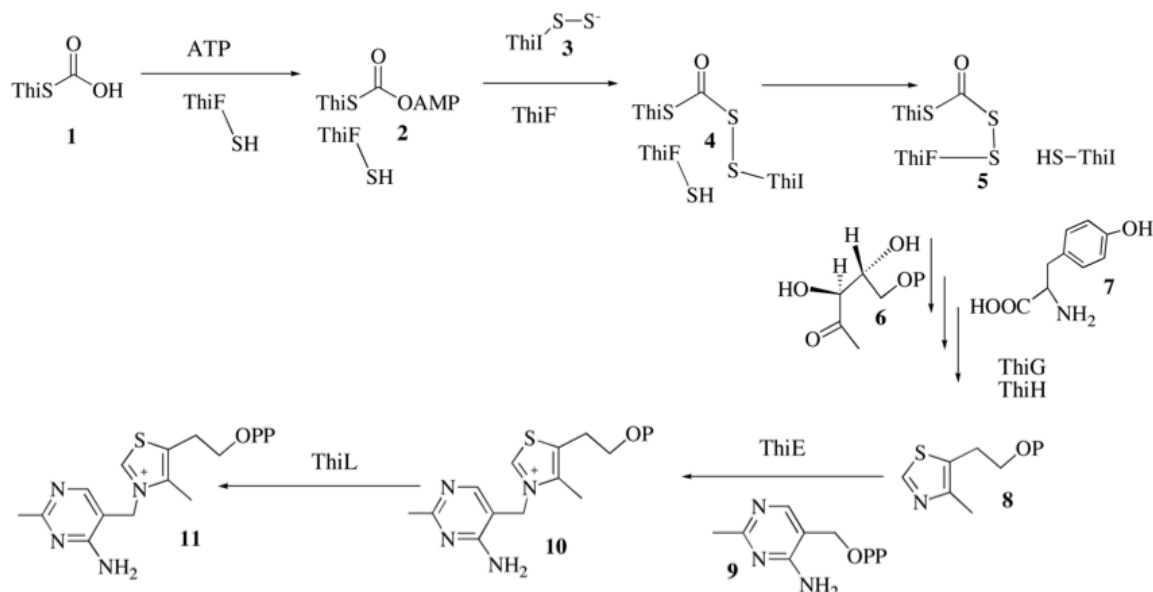


Figura 1.11. Esquema de la vía de síntesis de tiamina en *E. coli*. Se muestra como ThiF participa en la adenilación de la proteína ThiS de forma dependiente de ATP. Posteriormente es capaz de formar un puente persulfuro con esta misma a través de dos residuos de cisteína, siendo este uno de los pasos en la síntesis de tiamina en bacterias (Lehmann *et al.*, 2006).

ii) Proteína MoeB

De manera similar a ThiF, la proteína homóloga de TcdA, MoeB, es una proteína de tipo E1 capaz de adenilar MoaD, proteína de tipo ubiquitina. En este caso estas proteínas se encuentran en la ruta de síntesis del cofactor molibdenum (MoCo) (Lake *et al.*, 2001). MoeB es la encargada de adenilar el dominio C-terminal de MoaD, lo que posteriormente pasa a convertirse en un tiocarboxilato, mediante la acción de una sulfurotransferasa (Lake *et al.*, 2001). Estas reacciones corresponden a los primeros pasos en la síntesis del cofactor molibdenum, cofactor empleado por enzimas como la sulfito oxidasa o la aldehído oxidasa.

Por lo tanto se puede ver que el par MoeB-MoaD y ThiF-ThiS tienen funciones similares, en cuanto a la capacidad de adenilación de moléculas de tipo ubiquitina. Esto sugiere la existencia de un origen común de los sistemas E1/ubiquitina (eucariotas) y ThiF/ThiS y MoeB/MoeD en bacterias (Burrough *et al.*, 2009). En la figura 1.12. se muestra el plegamiento típico de esta superfamilia.

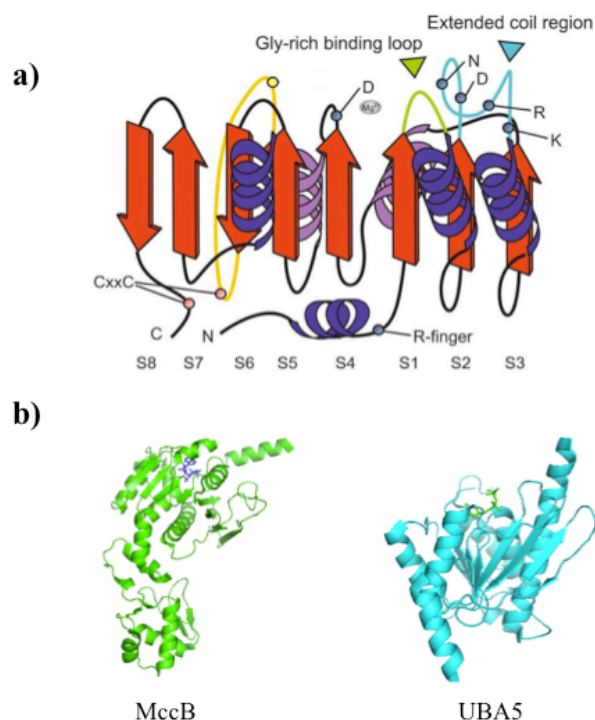


Figura 1.12. Diagrama topológico representativo de la estructura tridimensional de las proteínas E1 o de tipo E1 y ejemplo de estructuras resueltas.

a) Se muestran los diferentes componentes que caracterizan a esta superfamilia, observándose el dedo de arginina (*R-finger*), el dominio rico en glicinas (*Gly-rich binding loop*), así como el dominio de unión a zinc (*CxxC*). Las letras D, N, R y K, se refieren a los aminoácidos aspartato, asparagina, arginina y lisina respectivamente en *extended coil region*. b) Estructuras de los monómeros de las proteínas E1 MccB (pdb 3h9j) y UBA5 (pdb 3h8v) (imagen modificada de Burroughs et al., 2009). S, se refiere a *strand* (hebras) y H a *helix* (hélice).

En cuanto al dominio de tipo E1, estructuralmente hablando, éste posee tres capas de tipo α/β (Burroughs *et al.*, 2009) (figura 1.12). El núcleo de todo el plegamiento es de tipo Rossmann, típico de unión a nucleótidos. Además, las proteínas de tipo E1 poseen un conjunto de hélices característico en el extremo N-terminal. Normalmente este dominio presenta una arginina que se proyecta hacia el centro activo del monómero vecino, ya que estas proteínas suelen ser homodímeros. Este residuo de arginina además es capaz de estabilizar las cargas negativas del fosfato durante su transferencia en el proceso de adenilación, por lo que esta zona se ha denominado como el *arginine finger* (o *R finger*) por algunos autores o dedo de arginina (Burrough *et al.*, 2009). Otra de las características de esta superfamilia es la existencia de un *loop* entre S2 y H2 (figura 1.12.a), donde encontramos residuos polares, fuertemente conservados, siendo esenciales para el proceso de adenilación (*extended coil region*). Además, entre S5 y S6 se encuentra el *crossover loop*, donde se haya una cisteína conservada responsable en las reacciones de tiolación. Otra característica de estas enzimas es la aparición de residuos de aspartato en S4, encargados de coordinar el átomo de Mg^{2+} que estabiliza la unión de la molécula de ATP al centro activo. La familia de proteínas de tipo E1 también se caracteriza por la presencia del motivo de secuencia de unión a zinc (*CxxC*) (Burroughs

et al., 2009), localizado en el dominio C-terminal de ThiF y MoeB, no conservado en TcdA.

1.4.3.3. Adenilación

El proceso de adenilación o AMPilación (*AMPylation*) es una modificación post-traducciona que consiste en la cesión del grupo -AMP de la molécula de ATP, a una proteína diana a través de un enlace covalente. El mecanismo enzimático de adenilación puede compararse al de las enzimas quinasas, en cuanto que ambas son capaces de llevar a cabo la hidrólisis de la molécula de ATP, a la vez que transportan el metabolito generado a una proteína diana (Woolary *et al.*, 2010). Este proceso es realizado, entre otras, por las enzimas E1 (o de tipo E1) que catalizan la primera reacción de activación (o de tipo ubiquitina). Esta primera reacción de adenilación va seguida de dos reacciones más, tioesterificación y transtioesterificación (Lorenz *et al.*, 2013).

En las proteínas bacterianas, ThiF y MoeB, el proceso de adenilación (o cesión de AMP) ocurre desde el dominio N-terminal de estas enzimas, previamente cargadas con ATP y Mg^{2+} , sobre el C-terminal de las enzimas de tipo ubiquitina, siendo ThiS en el caso de ThiF y MoaD en el caso de MoeB (Duda *et al.*, 2005). En el caso del sistema perteneciente a la enzima ThiF, sobre la proteína sustrato (ThiS) se forma un grupo aciladenilato (ThiS-CO-AMP) que es necesario para la síntesis de tiamina.

En la zona N-terminal de ThiF y MoeB presenta plegamiento de tipo Rossmann y contienen el motivo de unión a nucleótidos llamado “*P-loop*” o lazo-P con la secuencia consenso (GXGXXG) (figuras 1.12 y 1.13). Este tipo de secuencia motivo se encuentra conservado en muchas enzimas de tipo E1.

<i>Ec</i>	TcdA	³⁷ GIGGVG ⁴²
<i>Ec</i>	ThiF	³⁵ GLGGLG ⁴⁰
<i>Ec</i>	MoeB	³⁸ GLGGLG ⁴³

Figura 1.13. Secuencias de unión de nucleótidos de las proteínas de tipo E1 de *E. coli* TcdA, ThiF y MoeB. Se muestra una zona del dominio N-terminal de las tres proteínas de tipo E1 de esta bacteria, donde se ve conservada la secuencia típica de unión a nucleótidos. Al menos en los casos de ThiF y MoeB la molécula de ATP es capaz de unirse específicamente a este dominio. Los números indican la posición del residuo en la secuencia aminoacídica de cada proteína.

En esta zona se produce la unión de la molécula de ATP, formando parte del plegamiento de tipo Rossmann y a través del cuál se da la adenilación, al menos en los casos de ThiF y MoeB. Para completar esta zona de unión del ATP, además hacen falta residuos de los dos monómeros.

En el caso de la proteína ThiF, la molécula de ATP se encuentra en una cavidad hidrofóbica, donde encontramos los residuos de Ile34, Arg106, Leu107, Thr126 y Thr131 que interaccionan con la molécula de ATP. Otros residuos descritos son el Asp59 y Asp61 los cuales participan en las interacciones con los grupos 2'-OH y el 3'-OH del anillo de ribosa de la molécula de ATP, formando parte del motivo Dx₂D, conservado tanto en el caso de MoeB como de TcdA. La Lys83 también interacciona con el 3'-OH de la ribosa de la misma molécula. Por su parte los residuos Gly38, Arg70 y Asp127 estabilizan el fosfato α del ATP, mientras que el fosfato β interacciona con la Arg70, la Gln71 y la Lys83. La Arg70 establece conexión con el fosfato γ (Duda *et al.*, 2005), siendo este último residuo el que le da nombre al dominio *arginine finger* mencionado anteriormente, la cual se conserva tanto en el caso de MoeB como en el de TcdA. En cuanto al catión de magnesio que acompaña a la molécula de ATP, éste es coordinado por el Asp127 de ThiF, conservado de nuevo en las otras dos proteínas de tipo E1 bacterianas, TcdA y MoeB.

Otro de los residuos característicos de ThiF es la Cys184, a través de la cual forma el puente persulfido posteriormente con ThiS en la segunda reacción en la ruta de la síntesis de tiamina (figura 1.11). Este residuo se encuentra en MoeB (Cys187), pero no en TcdA, donde aparece una alanina.

Por su parte, en el caso de MoeB, los residuos implicados en la unión de la molécula de ATP están altamente conservados respecto a ThiF. Así la Arg73 interacciona con los fosfatos α y β , equivalente a la Arg70 de ThiF. La Lys86 de MoeB interacciona con el fosfato β (la cual se observa también en ThiF) y la Ser69, así como la Asn70 con el fosfato γ , residuos también presentes en el caso de ThiF. Asp130 coordina la molécula de Mg²⁺ (Lake *et al.*, 2001), equivalente al residuo Asp127 de ThiF. Se comprueba que los residuos implicados en el proceso de adenilación se encuentran conservados tanto en el caso de ThiF como de MoeB.

1.4.3.4. Actividad TcdA

La enzima TcdA es responsable de ciclar la cadena lateral de treonilcarbamoiladenosina de la base 37 de los ARNts que reconocen codones que comienza con adenosina (ANN), convirtiendo t⁶A37 en ct⁶A37, en un proceso dependiente de ATP (Miyachi *et al.*, 2013). En la figura 1.14 se muestran las diferentes

Introducción

partes que componen una molécula de ARNt, observándose como la base 37 se encuentra adyacente al anticodón.

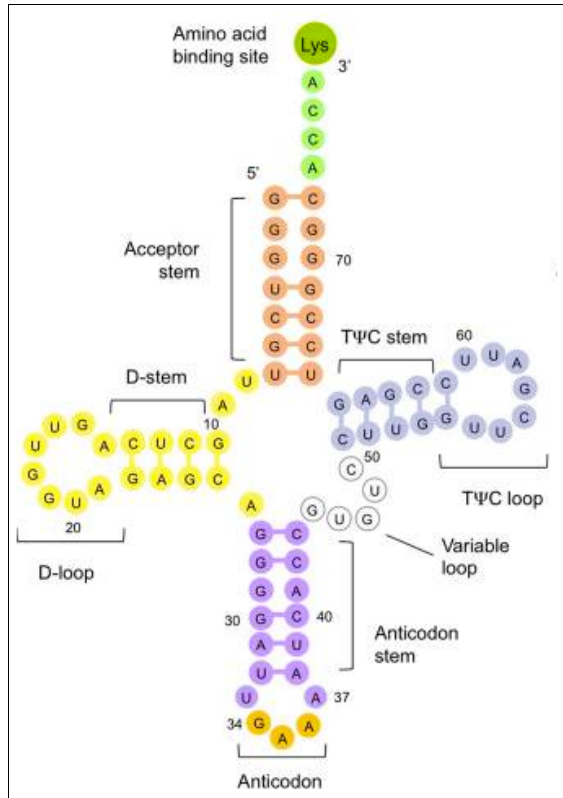


Figura 1.14 Figura esquemática de la estructura típica de trébol de un ARNt. Se muestran las partes principales de una molécula de ARNt. Las posiciones 34, 35 y 36 forman el conocido ASL (*anticodon stem-loop*), siendo la posición 34 la responsable en parte del fenómeno de degeneración del código genético. La base 37 es la diana de la enzima TcdA en aquellos ARNt que reconoce codones de tipo ANN (López-Esteva *et al.*, 2015).

1.4.3.5. Modificaciones sobre moléculas de ARNt

Se han descrito al menos 90 modificaciones que pueden sufrir los ARNt, habiendo en algunos organismos más información genética destinada a la modificación de estas moléculas que para el propio ARNt en sí (Björk GR, 1995). Este tipo de modificaciones en las moléculas de ARNt, en general, cumplen funciones tales como la estabilización de su estructura secundaria o la correcta lectura del ARN mensajero, por lo tanto es un paso esencial para el proceso de decodificación en cualquier organismo. Por ejemplo, las modificaciones en la posición 34 del anticodón (primera posición) son determinantes en la interacción anticodón-codón, lo que evita una lectura errónea (El Yacoubi *et al.*, 2012).

De las modificaciones descritas hasta el momento sobre moléculas de ARN de transferencia, las mejores estudiadas y comprendidas son aquellas llevadas a cabo sobre el ASL (*anticodon stem loop*), sobre todo en las posiciones 34 y 37 de la secuencia de

tRNA (figura 1.14). Esta posición 34 puede balancearse y aparearse con diferentes nucleótidos en la tercera posición del codón del ARN mensajero a través de una interacción de tipo no-Watson-Crick (Grabiell Torres *et al.*, 2014), lo que permite la degeneración del código genético. De hecho en la posición 34 se han observado múltiples modificaciones que van desde las hidroxilaciones, pasando por las metilaciones, tiolaciones, hasta las modificaciones de la adenina para producir inosina (El Yacoubi *et al.*, 2012), lo que indica la gran versatilidad que puede tener a la hora de decodificar la información genética. Estudiando este tipo de modificaciones y viendo la gran repercusión que tienen sobre la síntesis proteica, es sencillo pensar que el defecto en esta maquinaria puede conducir a la generación de enfermedades. Por ejemplo, en humanos, defectos en el gen codificante para la proteína FtsJ, una ARNt metiltransferasa, la cual actúa sobre la posición 34 de los ARNt para leucina, fenilalanina y triptófano, han sido relacionados con desórdenes mentales (Freude *et al.*, 2004). También se han registrado casos que relacionan defectos en enzimas que modifican la base 37 con enfermedades. Es el caso de la ARNt metiltransferasa de humanos TRMT12, encargada de modificar esta base para transformarla en wybutosina en el caso de ARNt para fenilalanina (Gabriel Torres *et al.*, 2014). Este gen se ha visto sobre-expresado en la mayoría de los casos de cáncer de mama estudiados (Rodríguez *et al.*, 2007). Defectos en estas modificaciones se han observado incluso en enfermedades tan comunes como la diabetes de tipo 2. De esta forma, defectos en el gen codificante para CDKAL1 en ratones, la enzima encargada de modificar la base 37 del ARNt para lisina, se han relacionado con fallos en la producción de diferentes proteínas, entre ella la proinsulina (Wei *et al.*, 2011).

Por último para entender la importancia de estas posiciones, se puede destacar que entre genomas de organismos distintos, codones sinónimos no se emplean con la misma frecuencia, hecho que se da gracias a este tipo de modificaciones (Gustilo *et al.*, 2008). Otro dato es que, ante el uso de codones raros, la estructura tridimensional de algunas proteínas puede variar respecto a si se hubiesen usado codones no raros (Kimchi-Sarfaty *et al.*, 2007), por lo que la influencia de este tipo de modificaciones puede ser de gran relevancia incluso a nivel de plegamiento proteico, colaborando en la diferenciación de distintos codones. Por ejemplo, en el caso de *E. coli* existen 6 codones para arginina, los cuales se encuentran reconocidos por cinco ARNt diferentes. En concreto, el ARNt^{Arg4}(UCU) es capaz de reconocer tan solo los codones AGA y AGG de los seis. Esto se explica gracias a las modificaciones que sufre la molécula de ARN, encontrándose una 2-tiocitidina en la posición 32, un 5-metilaminometiluridina (mnm⁵) en la posición 34 y la conocida t⁶A en la posición 37 (Sprinzl *et al.*, 1998). Por lo tanto, este tipo de modificaciones, aunque puedan parecer sutiles demuestran una gran importancia y precisión en el reconocimiento codón-anticodón.

1.4.3.6. Modificaciones t⁶A y ct⁶A en la posición 37

N⁶-treonilcarbamoiladenosina (t⁶A) es una modificación universal en la base 37 del ASL de varios ARNts, responsable del reconocimiento de los codones que comienzan con adenosina (ANN), a excepción de ARNt para metionina (El Yacoubi *et al.*, 2012; Thiaville *et al.*, 2014). Esta modificación se ha observado en todas las ramas de la vida, siendo descrita por primera vez en 1969 (Schweizer *et al.*, 1969) para el ARNt capaz de cargar el aminoácido triptófano.

La posición 37 de los ARNt se localiza junto a la tercera posición del anticodón. Esta posición cumple la función de estabilizar el apareamiento anticodón-codón mediante *stacking* de las bases en el sitio de decodificación del ribosoma, entre otras (figura 1.15) (Matuszewsiki y Sochacka, 2014), lo que la convierte en imprescindible en el proceso de decodificación y por tanto en la biosíntesis de proteínas.

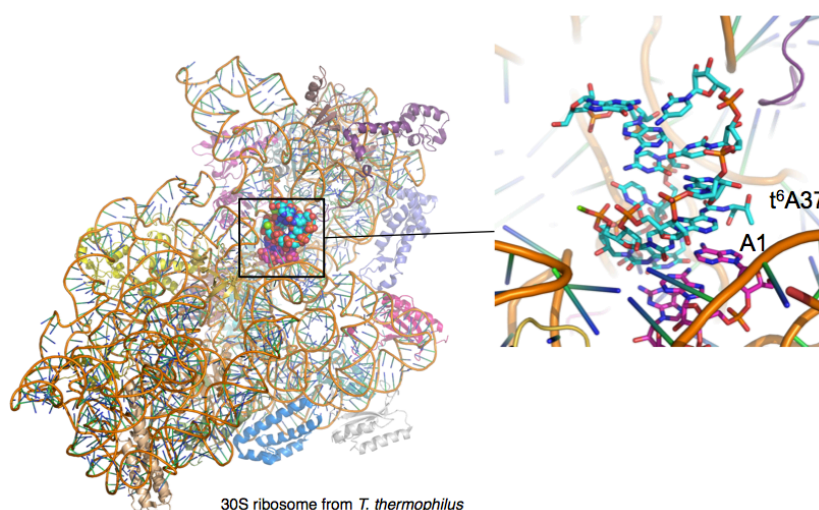


Figura 1.15. Estructura del t⁶A37 en el ASL del ARNt^{Lys}(UUU). En la parte izquierda de la imagen se puede observar el complejo formado por el AAA-ARNm, el ribosoma (subunidad 30S) de *Thermus thermophilus* (pdb 1xmo) y el ARNt^{Lys}(UUU). En la parte derecha se muestra la interacción mediante *stacking* del ARNt (con los átomos de carbono en cian) a través de la base t⁶A37, con el AAA-ARNm (átomos de carbono en violeta) mediante la base A1 del ASL (López-Esteva *et al.*, 2015).

En *E. coli* se dan 13 especies de ARNt que contienen la base t⁶A, o si no alguna de sus tres variantes: i) m⁶t⁶A (N6-metil-N6-treonylcarbamoiladenosina), ii) ct⁶A (la forma cíclica, poseyendo un anillo de oxazolano), y iii) ms²t⁶A (2-metiltio-N⁶-treonylcarba-

moyladenosina) (Machnicka *et al.*, 2013). La base de purina t⁶A37 se encuentra altamente conservada en los tres dominios de la vida, encontrándose modificada en la mayoría de casos (Helm y Alfonzo, 2014). Entre las funciones del nucleósido t⁶A se han descrito: i) la de permitir una mayor eficiencia en la aminoacilación de las moléculas de ARNt, ii) la unión del ARNt al sitio A del ribosoma, iii) la eficiencia en la translocación, iv) el mantenimiento del marco de lectura, así como v) una correcta localización de los codones de iniciación y de terminación (Matuszewsiki y Sochacka, 2014). Además, los impedimentos estéricos introducidos por la presencia de esta base, t⁶A37, previene entre otros el apareamiento con las bases número 33 y 38 del ARNt (Thiaville *et al.*, 2014), hecho que permite dar forma de “U” al *loop* anticodón (Sundaram *et al.*, 2000), esencial para cumplir su función en la traducción proteica, por lo que también cumple funciones estructurales.

Las conexiones moleculares formadas entre el átomo N1 de la adenina de la base de t⁶A37 y el átomo N11 del motivo treonina permiten que esta base se apile sobre la A38 en el *loop* anticodón y con la primera base (una adenina, A1) del codón, lo que lleva a pensar de nuevo en el papel estabilizador de esta base (t⁶A) en la interacción codón-anticodón (Murphy *et al.*, 2004).

Recientemente se ha descrito la hipermodificación de t⁶A37 consistente en la ciclación ct⁶A, llevada a cabo por TcdA (Miyauchi *et al.*, 2013). En dicho trabajo incluso se concluye que la modificación ct⁶A en la posición 37 es la forma real en la célula, siendo t⁶A un simple artefacto derivado de la metodología empleada en el aislamiento de este tipo de moléculas en organismos como *E. coli* (Miyauchi *et al.*, 2013). Esta afirmación se apoya en el hecho observado de que esta base ct⁶A parece ayudar al ARNt^{Lys} a decodificar codones como pueden ser AGA y UAG, con mayor eficacia que en el caso de t⁶A. Es de destacar, que todos los ARNt^{Lys} conocidos poseen la posición 37 modificada, ya sea o como t⁶A o como ms²t⁶A, aunque precisa de la modificación mnm⁵U34 para poder posicionar de forma correcta el ARNt en el ribosoma (Murphy *et al.*, 2004) y no producir fallos en el proceso de traducción. Esta mayor eficacia en la decodificación de la forma cíclica refuerza la hipótesis de Miyauchi *et al.*, 2013 de que se trata de una modificación real *in vivo*.

1.5. Interactores de CsdE

Como se ha comentado en párrafos anteriores, el sistema CSD era el menos conocido de todos aquellos descritos en la bacteria *E. coli*, en cuanto a la captación de átomos de azufre a través de cisteína desulfurasas. De esta forma en 2009 un estudio

Introducción

realizado por Trotter y colaboradores (Trotter *et al.*, 2009) identificaron las proteínas que podrían seguir la ruta mencionada anteriormente de transporte de azufre a través de las proteínas descritas para el sistema CSD. En la figura 1.16 se muestran los candidatos que surgieron como posibles interactores de CsdE mediante un experimento de co-elución. Las proteínas detectadas fueron entre otras: i) la proteína CsdA (la cisteína desulfurasa del sistema CSD), ii) HSP70 (una chaperona), ii) YgdL (o TcdA como se conoce gracias a su actividad enzimática) y iv) PrsA (una PRPP sintetasa).

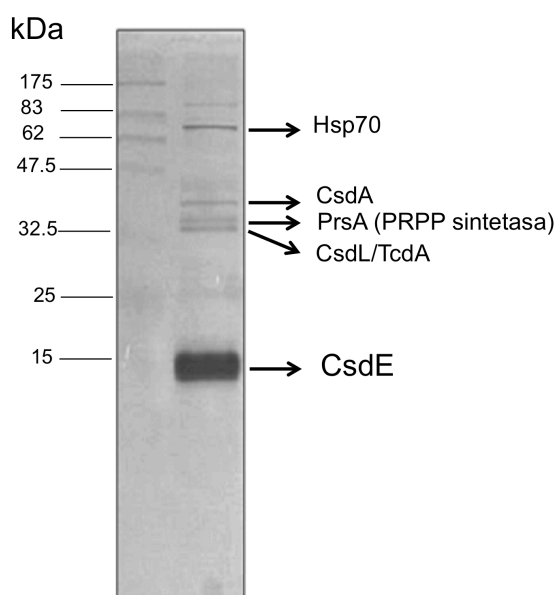


Figura 1.16. Experimento de coelución para la detección de posibles interactores de CsdE. La proteína CsdE (sobreexpresada) de *E. coli* se enganchó a una columna, a través de la cual fluyó el lisado de un cultivo de la misma bacteria. Posteriormente se eluye CsdE junto con las proteínas unidas a ella de forma específica en condiciones nativas, analizando posteriormente mediante PAGE-SDS y MALDI-TOF estas proteínas, con el fin de identificarlas (imagen modificada de Trotter *et al.*, 2009).

1.5.1. Fosforribosil transferasa o ribosa fosfato pirofosfoquinasa (PrsA)

La enzima PrsA, o PRPP sintetasa de *E. coli*, también es conocida como ribosa-fosfato pirofosfoquinasa (EC: 2.7.6.1). La reacción catalizada por PrsA tiene como sustratos las moléculas de ATP y ribosa 5-fosfato (R5P) para dar como producto AMP y PRPP (fosforribosil pirofosfato) (figura 1.17). De esta forma, esta actividad enzimática enlaza con:

- i) la ruta de la pentosa fosfato, pues recibe uno de sus productos finales (R5P). La ruta de las pentosas fosfato se puede dividir en dos fases. En la primera de ellas, conocida como la fase irreversible oxidativa, la glucosa una vez fosforilada es convertida a ribulosa 5-fosfato mediante una descarboxilación oxidativa, lo que genera NADPH. En la siguiente fase, la fase reversible no

oxidativa, la ribulosa 5-fosfato se puede metabolizar para originar ribosa 5-fosfato o xilulosa 5-fosfato.

- ii) la síntesis de purinas, las cuales se originan a partir de PRPP combinando diferentes sustratos. La molécula de PRPP en combinación con adenina origina AMP, mientras que con guanina dará GMP. Por su parte, gracias a la reacción entre hipoxantina y PRPP se obtiene IMP. Todas estas bases pueden formar tanto parte del ARN como del ADN.
- iii) la ruta de las pirimidinas se inicia con la reacción llevada a cabo por la OPRT (Orotato Fosforribosil Transferasa), donde se toman como sustrato PRPP y ácido orótico (OA) para originar OMP, el cual puede dar UMP mediante su descarboxilación.
- iv) la síntesis de NAD (Nicotinamida Adenina Dinucleótido) y NADP (Nicotinamida Adenina Dinucleótido Fosfato) (Breda *et al.*, 2002).

Del mismo modo, enlaza con la primera reacción de la ruta que lleva a cabo la síntesis de histidina, la cual comienza con PRPP y ATP (Kulis-Horn *et al.*, 2014), así como con la síntesis de triptófano en Bacteria, ruta que necesita de PRPP (Hove-Jensen, 1988; Breda *et al.*, 2002; Yanofsky, 2007).

1.5.2. Actividad PRPP sintetasa

Siguiendo con la reacción llevada a cabo por la PRPP sintetasa, ésta cataliza la transferencia del grupo β - γ -pirofosforil desde una molécula de ATP a otra de ribosa 5-fosfato, siendo estos dos sus sustratos (figura 1.17). Esta reacción es dependiente de Mg^{2+} por su efecto quelante sobre la molécula de ATP, además de promover la actividad de la enzima (White *et al.*, 1971). Esta reacción da lugar a la formación de una molécula de PRPP (5-fosfo-D-ribosil α -1-difosfato) y otra de AMP (Bower *et al.*, 1989).

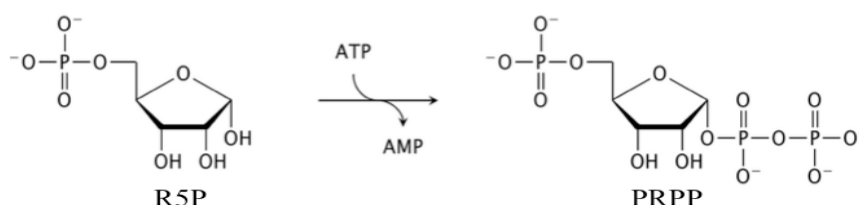


Figura 1.17. Reacción llevada a cabo por la PRPP sintetasa. Transformación de R5P a PRPP gracias a la donación de los fosfatos β y γ del ATP, originando una molécula de AMP (Breda *et al.*, 2012).

Además de ser clave en muchas rutas metabólicas, el PRPP es además precursor de componentes de la pared celular en bacterias tales como micobacteria (Scherman *et al.*, 1996). Su intervención en este tipo de vías metabólicas lo convierte en un buen candidato de estudio en cuanto al desarrollo de fármacos, siendo vías que intervenidas pueden afectar en gran medida la viabilidad celular de patógenos. Además se ha demostrado que el carbono de estos componentes pueden retornar a la vía de las pentosas fosfato directamente, por lo que indirectamente puede afectar a procesos tan importantes como la glucólisis (Switzer *et al.*, 1969) en las células patógenas.

1.5.3. Clasificación de PRPP sintetasas

Esta actividad PRPP sintetasa ha sido descrita en todas las ramas de la vida. Las enzimas capaces de llevar a cabo esta actividad han sido clasificadas dentro de tres clases en función de sus propiedades enzimáticas.

i) La primera de ellas, la clase I, incluye las proteínas procedentes de organismos como *E. coli*, *Salmonella typhimurium*, *Bacillus subtilis* o las de mamíferos. En todas ellas se describe una estructura cuaternaria de homohexámero, dándose la inhibición de su actividad alostéricamente por medio de ribonucleósidos de purina difosfato, como ADP o GDP. Además todas ellas requieren de fosfato inorgánico (Pi) para poder llevar a cabo su actividad de forma eficiente, además de tener una función estabilizadora del complejo (Breda *et al.*, 2012; Willemoës *et al.*, 2000; Switzer *et al.*, 1969).

ii) Las enzimas de clase II se encuentran únicamente en plantas, donde no necesitan del Pi para llevar a cabo su actividad enzimática, además de no mostrar inhibición por moléculas como ADP o GDP. En cuanto al donador de los átomos de fosfato, en este caso además de tomarlos de la molécula de ATP, también se han observado como sustratos moléculas como GTP, CTP y UTP. La estructura cuaternaria descrita en este caso ha sido de tetrámero (Krath y Hove-Jensen, 2001).

iii) Las PRPP sintetasas de clase III han sido descritas en la arquea *Methanocaldococcus jannaschii*. En este caso la enzima necesita de Pi para llevar a cabo su función y el único donador descrito de fosfatos ha sido el ATP. Además el ADP no actúa en este caso como inhibidor. Otra diferencia es su estructura cuaternaria, siendo en este caso un tetrámero (Kadziola *et al.*, 2005; Krath y Hove-Jensen, 2001). Otro ejemplo de la clase III es la PRPP sintetasa codificada en el genoma de *Thermoplasma volcanium*, poseyendo en este caso una estructura cuaternaria de homodímero. Tampoco ha sido descrito aquí ningún sitio de unión a ADP, por lo que no es inhibida por esta molécula, hecho por el cual se cataloga dentro de la clase III (Cherney *et al.*, 2011).

1.5.4. Estructura cuaternaria de las enzimas PRPP sintetasas

En cuanto al estado oligomérico de estas proteínas se han descrito como homohexameros las de clase I, como es el caso de las PRPP sintetasas de *E. coli* y *B. subtilis*. Sin embargo, en el caso de *Mycobacterium tuberculosis* se han observado la existencia de diferentes estados oligoméricos, atendiendo al tipo de molécula que ocupe el centro activo. De esta forma se ha descrito como un hexámero en su forma apo o unida a ribosa 5-fosfato, como un tetrámero cuando se encuentra unida a ADP y como un dímero cuando ha unido ATP (Breda *et al.*, 2012). En cuanto a los residuos implicados en su actividad, las de la clase I son las mejores caracterizadas, en concreto la PRPP sintetasa de *B. subtilis*, la cual guarda un porcentaje de identidad del 50% con la enzima de *E. coli*.

<i>Ec</i> PRPP sintetasa	¹⁹⁰ AIID KRR PRAN ²⁰⁰
<i>Bs</i> PRPP sintetasa	¹⁹³ AIID KRR PRPN ²⁰³

Figura 1.18. Similaridad de secuencia de las PRPP sintetasas de *E. coli* y de *B. subtilis* en el centro activo. En ambas proteínas se conservan residuos descritos como críticos a la hora de llevar a cabo la reacción enzimática (Lys197 y Arg199 de la secuencia de *B. subtilis*, siendo los residuos 194 y 196 en el caso de *E. coli*), mostrándose un gran parecido en esta zona de la secuencia.

Los residuos detectados en la secuencia de la enzima de *B. subtilis* como esenciales para llevar a cabo su actividad son las Lys197 y la Arg199, correspondiendo a los residuos Lys194 y Arg196 de *E. coli*, respectivamente (figura 1.18) (Hove-Jensen *et al.*, 2005), y se hayan involucrados en la velocidad de la reacción.

Otro residuo importante detectado en la secuencia de la enzima en estos dos organismo es el Asp128 de *E. coli*, correspondiente al Asp134 en *B. subtilis*, responsable de la unión del átomo de Mg²⁺, necesario para la llevar a cabo la reacción de forma eficiente (Bower *et al.*, 1989).

1.5.4.1. La enzima PRPP sintetasa como diana terapéutica

El estudio de la actividad enzimática de PRPP sintetasa abre nuevas vías de terapias en medicina. Por un lado, al estar relacionada directamente con la vía de síntesis *de novo* de purinas y pirimidinas, está directamente relacionada con la duplicación celular así

como con la síntesis de material genético como ARN o ADN. En concreto, se ha visto en líneas celulares humanas de carcinoma como durante el proceso de duplicación celular se da un aumento de la concentración de las pentosas fosfato (Fridman *et al.*, 2013). Además se ha observado que cuando el oncogen *myc* humano se ve sobreexpresado en ciertos tumores, afecta de forma positiva la producción de la PRPP sintetasa, hecho concordante con el gran aporte de purinas y pirimidinas que necesita una célula maligna para dividirse rápidamente (Cunningham *et al.*, 2014). Por tanto, el estudio de moléculas dirigidas a inhibir esta enzima en células capaces de producir tumores podría dar lugar al descubrimiento de nuevos tratamientos en determinados tipos de cáncer. Defectos en PRPP sintetasa en humanos, también se han asociado con diversas enfermedades. Por ejemplo, la sobreexpresión de PRPP sintetasa puede conducir a la acumulación de ácido úrico y por tanto al desarrollo de gota o anomalías neuronales como una consecuencia del defecto de esta actividad (Li *et al.*, 2007). Estos hechos destacan la importancia del buen funcionamiento de este tipo de vías en la fisiología humana.

En bacterias el papel central de la PRPP sintasa en mantener el aporte de nucleótidos necesarios para la división celular sugiere que las PRPP sintetasas de patógenos podría ser también una diana farmacológica para el desarrollo de antibióticos. De esta forma, el estudio estructural y cinético de proteínas de patógenos u organismos modelo es de gran interés para el tratamiento de diferentes enfermedades.

1.5.4.2. Estructuras resueltas de PRPP sintetasas

Hasta la fecha se conocen las estructuras tridimensionales de PRPP sintetasas de organismos como *E. coli*, *B. subtilis*, *Homo sapiens* o *Burkholderia pseudomallei*. En el caso de la enzima de *E. coli* la proteína ha sido cristalizada en su forma apo (simplemente con una molécula de Mg^{2+}) a una resolución de 2,71 Å (pdb 4s2u). La estructura de *B. subtilis* ha sido resuelta a una resolución de 2,80 Å, con las moléculas α -metilen adenosina monofosfato, Cd^{2+} y SO_4^{2-} (pdb 1ibs) (Eriksen *et al.*, 2002). Por último la enzima perteneciente a *H. sapiens* fue resuelta a una resolución de 2,20 Å, encontrando en el centro activo la molécula SO_4^{2-} (pdb 2h06) (Li *et al.*, 2007) En la figura 1.19 se muestran las estructuras.

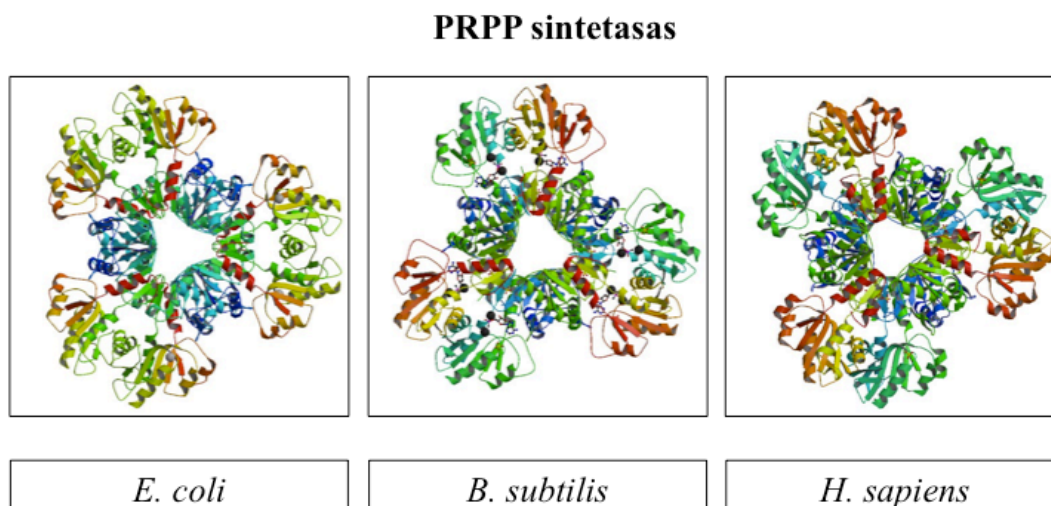


Figura 1.19. Estructuras resueltas de PRPP sintetetasas de diferentes organismos. Se muestran las estructuras de los organismos *E. coli* (pdb 4s2u), *B. subtilis* (pdb 1ibs) y *H. sapiens* (pdb 2h06).

1.5.5. Estudio de OPRT y su relación con la PRPP sintetasa: entrada en la ruta de síntesis *de novo* de pirimidinas

Para entender la función del sistema de CSD, sistema modelo de objeto de estudio en la presente tesis doctoral, se abordó el estudio de la proteína PRPP sintetasa, cuya función se halla estrechamente ligada a la función de OPRT (orotato fosforribosil transferasa) relacionada con el primer paso de la ruta de síntesis de pirimidinas.

En la ruta de biosíntesis *de novo* de pirimidinas uno de los destinos que sigue la molécula de PRPP sintetizada por la enzima PRPP sintetasa es la de ser sustrato de la enzima OPRT (orotato fosforribosiltransferasa) (EC: 2.4.2.10). OPRT cataliza la conversión de una molécula de PRPP y OA en orotidina 5'-monofosfato (OMP) más una molécula de pirofosfato (figura 1.20). Esta reacción necesita de la presencia de magnesio para poder llevarse a cabo.

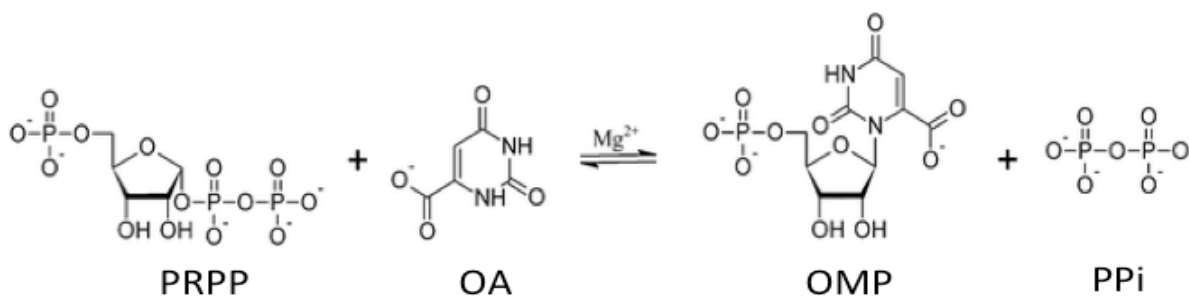


Figura 1.20. Reacción enzimática llevada a cabo por la OPRT de forma dependiente de magnesio. En la reacción se produce la fusión de la molécula de ácido orótico (OA), al grupo fosforribosil de la molécula de PRPP, liberándose el grupo pirofosfato de esta última molécula, para dar como resultado una molécula de OMP y otra de pirofosfato (PPi). La reacción es reversible (Breda *et al.*, 2012).

La formación de la molécula de ácido orótico ha implicado los primeros 4 pasos de la ruta en los que participa carbamoil fosfato sintetasa, aspartato transcarbamoilasa, dihidroorotasa, y dihidroorotato deshidrogenasa (fig. 1.21).

A través de la decarboxilación de la molécula de OMP se origina UMP que constituye la primera pirimidina dentro de la ruta de síntesis *de novo* de pirimidinas (figura 1.21) (Moffatt y Ashihara, 2002). A través de la molécula de UMP se puede llegar a todos sus derivados con distintos grados de fosforilación (UDP y UTP), así como a la molécula de CTP, gracias a la CTP sintetasa. Además, a través de la molécula de UMP se obtiene TMP, gracias a la acción de la timidilato sintetasa (Moffatt y Ashihara, 2002).

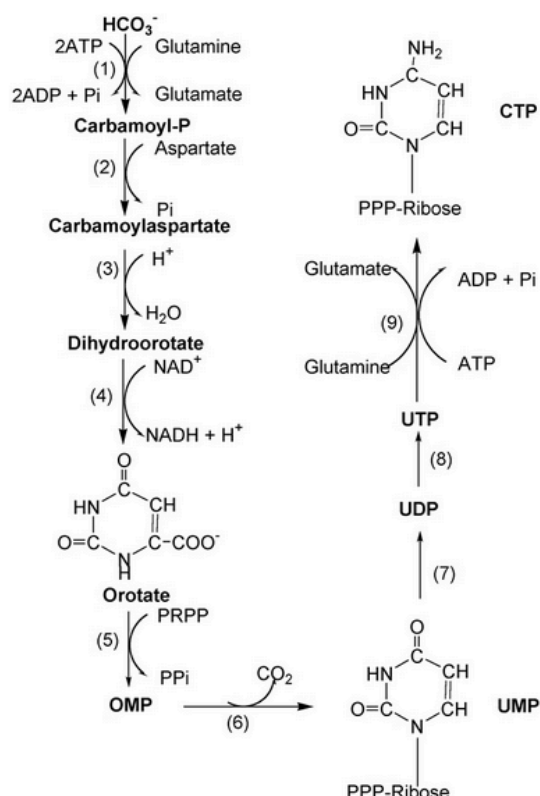


Figura 1.21. Ruta de biosíntesis *de novo* de pirimidinas. A partir de la ruta mostrada los organismos son capaces de sintetizar todas las pirimidinas. Las enzimas que intervienen son: 1) carbamoyl fosfato sintetasa, 2) aspartato transcarbamoilasa, 3) dihidroorotasa, 4) dihidroorotato deshidrogenasa, 5) orotato fosforribosil transferasa, 6) OMP descarboxilasa, 7) UMP quinasa, 8) nucleósido difosfato quinasa y 9) CTP sintetasa. A través de esta vía es posible obtener todas las pirimidinas y sus derivados (Moffatt y Ashihara, 2002).

La ruta de síntesis *de novo* de pirimidinas toma gran relevancia en el metabolismo celular, pues es la principal vía por la que se pueden obtener todas las pirimidinas para llevar a cabo diferentes procesos celulares, entre ellos el de división celular o síntesis de moléculas como ARN. Además esta ruta interviene también directamente en la síntesis de azúcares, otros nucleótidos, moléculas como fuente de energía y polisacáridos que formarán parte de la pared celular tanto de microorganismos como de plantas (Ong *et al.*, 2013; Schröder *et al.*, 2005).

La OPRT de *E. coli* pertenece a la familia I de las fosforribosiltransferasas (González Segura *et al.*, 2007), siendo característico de ella el plegamiento de tipo Rossmann (Grubmeyer *et al.*, 2012), que como ya se ha mencionado anteriormente, es un clásico motivo estructural de unión a nucleótidos.

En cuanto a la influencia de esta actividad en humanos, aquí no encontramos la actividad OPRT como una sola entidad proteica, sino que se encuentra en la enzima uridina monofosfato sintasa (EC: 4.1.1.23), la cual es capaz de llevar a cabo la actividad OPRT y la OMP descarboxilación (Wang *et al.*, 2007). En humanos, la carencia de esta actividad enzimática es hereditaria y puede generar desórdenes por un retraso en el crecimiento, además de causar una anemia grave y secreción de grandes cantidades de

ácido orótico en la orina (Kim *et al.*, 2007), por lo que de nuevo nos encontramos ante una vía metabólica de relevancia.

1.5.5.1. La enzima OPRT como diana terapéutica

Uno de los grandes intereses que posee esta actividad enzimática es su asociación con la ruta de síntesis *de novo* de las pirimidinas, por lo que a partir de ella se puede controlar fácilmente toda esta vía, así como las consecuencias de ello. Además se ha observado tanto en procariotas como en eucariotas que la falta total de pirimidinas conduce a la muerte por falta de timidina (Ong *et al.*, 2013), lo que demuestra que esta vía es esencial.

En concreto no se buscan inhibidores de esta actividad enzimática, sino moléculas que una vez entradas en la vía son metabolizadas, lo que las convierte en metabolitos perjudiciales para la viabilidad celular, en concreto de ciertos patógenos. Es el caso del análogo del ácido orótico (OA), uno de los sustratos de la OPRT, 5-FOA (*5-fluoroorotic acid*) el cual es metabolizado hasta 5-FUMP (*fluorouridine monophosphate*), incorporándose a las moléculas que precisen de UMP. Además se ha registrado que la molécula 5-FUMP es capaz de inhibir la actividad de la timidilato sintasa, impidiendo la síntesis de esta base esencial para la replicación del ADN. Estos efectos en concreto se han observado en el parásito humano *Trypanosoma brucei*, causante de la tripanosomiasis (Ong *et al.*, 2013). Por lo tanto, este tipo de moléculas actuarían como prodrogas, esto es, se activan tras ser metabolizadas, lo que puede ocasionar efectos adversos sobre el metabolismo celular (Wang *et al.*, 2007; López-Esteba *et al.*, 2011), mientras que si no sufren ninguna modificación química son totalmente inocuas para los organismos.

Otro caso de gran relevancia es el descubrimiento de que análogos de pirimidinas, como 5-FUMP, son capaces de detener en un 90% la replicación del parásito causante de la toxoplasmosis, *Toxoplasma gondii* (Javaid *et al.*, 1999). Algo similar ocurre en el caso de la especie de *Plasmodium falciparum*, causante de la malaria, el cual ve inhibido en parte su desarrollo ante la asimilación de moléculas tales como el 5-FOA. En este parásito toma mayor relevancia este tipo de moléculas, por su incapacidad de tomar pirimidinas exógenas, teniendo que sintetizarlas todas el propio parásito (Rathod *et al.*, 1989).

Todas estas observaciones convierten a esta vía de síntesis *de novo* en esencial en cuanto a su estudio y comprensión, con el fin de entender las diferencias entre la vía de patógenos y de mamífero y poder así desarrollar fármacos de mayor potencia y específicos de agentes patógenos.

1.5.5.2. Estructuras resueltas de OPRT

La enzima OPRT de *E. coli* fue resuelta a una resolución de 2,40 Å con una molécula de SO_4^{2-} en su centro activo (pdb 1oro) (Henriksen *et al.*, 1996). Por su parte la estructura de la enzima de *S. typhimurium* fue resuelta a una resolución de 2,30 Å con las moléculas Mg^{2+} , PRPP y OA en el centro activo (pdb 1opr) (Scapin *et al.*, 1995). La estructura de la enzima de *S. cerevisiae* fue resuelta en presencia de tres combinaciones diferentes de moléculas. En presencia de Mg^{2+} (pdb 2pry) a 2,35 Å ; con OMP (pdb 2prz) a 1,89 Å o con PRPP, OA y Mg^{2+} (pdb 2ps1) a 1,75 Å en su centro activo (González-Segura *et al.*, 2007). Por último, recientemente se ha resuelto la estructura de la OPRT de *P. falciparum* (pdb 4fym), a una resolución de 2,60 Å y con la molécula SO_4^{2-} en su centro activo (Kumar *et al.*, 2015). En la figura 1.22 se muestran las estructuras resueltas de las enzimas de estos organismos.

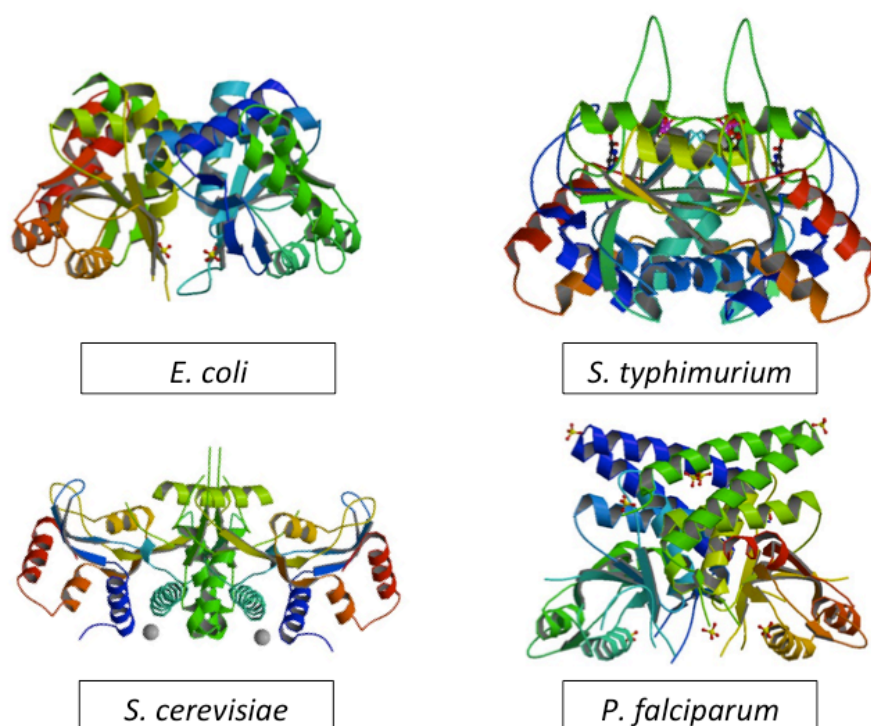


Figura 1.22. Estructuras resueltas de OPRT de diferentes organismos. Se muestran las estructuras de los organismos *E. coli* (pdb 1oro), *S. typhimurium* (pdb 1opr), *S. cerevisiae* (pdb 2ps1) y *P. falciparum* (pdb 4fym).

Objetivos

2. Objetivos

Cisteína desulfurasa persulfurada CsdA sola y en complejo con CsdE

1. Estudio estructural de la cisteína desulfurasa CsdA de *E. coli* en sus formas persulfurada y sulfatada.
2. Determinación estructural del complejo heterotetrámero de (CsdA-CsdE)₂ en su forma persulfurada.
3. Caracterización del transporte de S (fenómeno de *crosstalk*) desde componentes del sistema CSD (CsdA) al sistema Suf (SufE), ambos de *Escherichia coli*.

ARNt treonilcarbamoiladenosina deshidratasa, TcdA

4. Determinación estructural de la proteína TcdA de *Escherichia coli*, acomplexada con ATP-Mg²⁺ y AMP-Mg²⁺.
5. Caracterización de la interacción TcdA-ARNt^{Lys}(UUU).
6. Caracterización del modo de interacción de la proteína TcdA de *E. coli* con el sistema CSD del mismo organismo.

Fosforribosil pirofosfato sintetasa (PRPP sintetasa) y Orotato fosforribosil transferasa, OPRT

7. Estudio estructural de la PRPP sintetasa de *E. coli* en complejo con ATP
8. Determinación estructural y caracterización cinética de la proteína OPRT de *E. coli* acomplexada con los análogos de sustrato, 5-FOA y 5-AOA.
9. Determinación estructural y caracterización cinética de la proteína OPRT de *Mycobacterium tuberculosis*.

Materiales y Métodos

3. Materiales y Métodos

3.1. Herramientas bioinformáticas

3.1.1. Secuencias de ADN y proteínas

Con el fin de diseñar los experimentos de forma correcta y estudiar las proteínas de interés se hicieron búsquedas de los genes codificantes para las proteínas diana. Para ello, se emplearon las bases de datos del *National Center for Biotechnology Information* (NCBI, www.ncbi.nih.gov/pubmed) y del *Universal Protein Resource* (UNIPROT, www.uniprot.org). Dentro de esta última base de datos se empleó la herramienta *Basic Local Alignment Search Tool* (BLAST), la cual es capaz de encontrar regiones de similitud local entre secuencias de nucleótidos o de proteínas, al compararlas con las secuencias depositadas en la base de datos. También se usó la herramienta de alineamiento de secuencias de proteínas (*Align*), de la base de datos de UNIPROT, la cual permite comparar estas secuencias residuo a residuo.

En cuanto al estudio de las moléculas de ARN de transferencia (ARNt) de *E. coli* se empleó la base de datos *Genomic tRNA Database* (http://lowelab.ucsc.edu/GtRNAdb/Esch_coli_K12/Esch_coli_K12-by-isotype.html), con el fin de obtener las secuencias de los distintos ARNt empleados para su clonaje.

Por otro lado, en cuanto al diseño de las estrategias para el clonaje, expresión y purificación de las diferentes proteínas se empleó el portal ExPASy (*SIB Bioinformatics Resource Portal*, www.expasy.org), donde es posible calcular valores teóricos de parámetros químico-físicos de las proteínas como la masa molecular (*Molecular Weight*), el punto isoelectrico (pI), o el coeficiente de extinción molar, entre otros.

3.1.2. Estructura 3D: PDB (*Protein Data Bank*)

Con el fin de comparar las estructuras tridimensionales de las proteínas diana con otras estructuras resueltas, se empleó la base de datos PDB (<http://www.rcsb.org>). Ésta acumula un repertorio único de información sobre estructuras tridimensionales tanto de proteínas como de ácidos nucleicos, así como de diferentes complejos. Del mismo modo, una vez resueltas las estructuras de las proteínas diana, PDB es la base de datos donde se aportan las coordenadas de las proteínas resueltas.

3.2. Métodos generales en clonaje, expresión y purificación de proteínas recombinantes en *E. coli*

3.2.1. Clonaje

3.2.1.1. PCR (*Polymerase Chain Reaction*)

El clonaje de las proteínas recombinantes en todos los casos se realizó por la vía clásica de restricción-ligación. Una vez obtenida la secuencia para el gen que codifica para la proteína que se desea clonar de la base de datos NCBI, se lleva a cabo el diseño de los cebadores con el fin de amplificar el gen desde el ADN molde por la técnica de la PCR (*Polymerase Chain Reaction*), reacción en cadena de la polimerasa. Ésta técnica de biología molecular está basada en la capacidad de la ADN polimerasa de sintetizar una hebra de ADN complementario a la secuencia de la molécula molde de ADN, originando lo que se denominará de aquí en adelante el inserto. Además la ADN polimerasa necesita de la presencia de cebadores, a partir de los cuales pueda añadir el primer nucleótido. Estos cebadores son oligonucleótidos de secuencia corta que se unen de forma complementaria al ADN molde, actuando como punto de inicio de la síntesis de ADN. La reacción de PCR, además de la ADN polimerasa, los cebadores y el ADN molde, necesita de dNTPs (desoxinucleótido trifosfato), el sustrato de la reacción. En cuanto a la ADN polimerasa empleada, en este primer paso siempre fue la *Phusion® High-Fidelity DNA polymerase* (NEB, New England Biolabs).

Los oligonucleótidos empleados en cada caso llevan en los extremos de su secuencia dianas específicas para enzimas de restricción, con el fin de ligarlas en vectores de expresión para *E. coli*, empleado como sistema de expresión.

En cuanto al método seguido para llevar a cabo la reacción de PCR, una vez todos los componentes se mezclan, se realiza un primer paso de desnaturalización de las hebras del ADN molde a 370 K durante 5 minutos. Seguidamente se lleva a cabo un ciclo de 370 K durante 30 segundos más 330 K durante 30 segundos en el proceso de *annealing* del ADN molde con los cebadores y de 345 K durante el proceso de extensión. La duración de éste último depende de la longitud del ADN que se quiere conseguir, tomando como referencia que la ADN polimerasa necesita de 1 minuto aproximadamente para sintetizar una hebra de 1 kb (kilobase). El ciclo mencionado se repetirá 34 veces tras los cuales se lleva a cabo un último paso de extensión de 10 minutos. Todo este proceso es llevado a cabo gracias a un termociclador (Mastercycler® Nexus de Eppendorf).

El clonaje de los genes codificantes de los ARNt estudiados se llevó a cabo de forma similar al descrito para los genes de las proteínas, salvo que las secuencias de los genes fueron tomadas del portal *Genomic tRNA Database*.

Los genotipos de las cepas empleadas tanto para la obtención de genes, como para los procesos de clonaje y expresión de proteínas, se detallan sus genotipos en la tabla 3.1.

Cepa <i>E. coli</i>	Genotipo
XL1blue	<i>endA1 gyrA96(nal^R) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB⁺ lacI^q Δ(lacZ)M15] hsdR17(r_K m_K⁺)</i>
BL21(DE3)	<i>F⁻ ompT gal dcm lon hsdS_B(r_B⁻m_B⁻) λ (DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) [malB⁺]_{K-12} (λ^S)</i>

Tabla 3.1. Cepas de *E. coli* empleadas en esta tesis. La cepa XL1blue (*Stratagene*) se empleó para la obtención de los genes de *E. coli* (codificantes de las proteínas expresadas) y BL21(DE3) (*Novagen*) para la expresión de los genes codificados

3.2.1.2. Electroforesis de ADN

Una vez llevada a cabo la reacción de PCR, las muestras obtenidas son analizadas mediante una electroforesis de ADN en un gel de agarosa al 1%, empleando como solvente el tampón TBE (100 mM Tris, 100 mM ácido bórico y 2 mM EDTA (ácido etilendiaminotetraacético), a un pH final de 8,0). A la hora de gelificar la agarosa se añadió SYBR® Safe, el cual permite la visualización de ácidos nucleicos mediante el uso de luz ultravioleta. La solución empleada para correr el gel fue también TBE. El gel de agarosa se carga con 5 µL de los marcadores correspondientes en un pocillo (NZYDNA Ladder III) y 10 µL de cada muestra obtenida de la PCR más 2 µL del tampón de carga (6x) en los diferentes pocillos. El tampón de carga en este caso se compone de glicerol al 30% (v/v) y azul de bromofenol.

La electroforesis se llevó a cabo en una cubeta para tal fin durante 25 minutos a 150 V a temperatura ambiente, visualizando el gel posteriormente en un G:Box F3 equipado con un transiluminador de UV y un filtro de BrEt (*Syngene*).

3.2.1.3. Digestión y ligación

Tras la comprobación mediante electroforesis de que el inserto obtenido del proceso de PCR tiene el tamaño adecuado, se digieren los extremos de éste con las enzimas de restricción para las que fueron diseñados los oligonucleótidos. Las enzimas de restricción empleadas corresponden a la casa NEB y son distintas en los dos oligonucleótidos empleados por PCR. El mismo proceso se lleva a cabo con el vector de expresión elegido para clonar el inserto obtenido. Sin embargo, una vez digerido el vector con las mismas enzimas de restricción que el inserto, es necesario además desfosforilar los extremos de éste mediante el empleo de la fosfatasa antártica de NEB. A continuación se purifican tanto el inserto como el vector digerido mediante el uso de mini-columnas de la casa QIAGEN que unirán el ADN diana, eliminando la ADN polimerasa, así como las enzimas de restricción, dejando tanto el inserto como el vector en tampón TE (10 mM Tris pH 8,0 y 1 mM EDTA) listo para su uso.

Una vez purificados estos componentes se realiza la reacción de ligación, la cual implica el uso de los fragmentos digeridos anteriormente (inserto y vector), en un ratio molar 3:1, aplicando 100 ng de vector y la correspondiente cantidad de inserto, según la masa molecular de éste. Además debe aplicarse la ligasa de ADN T4 (NEB) en su tampón correspondiente. Una vez mezclados los componentes, la reacción es llevada a cabo a 295 K durante 30 minutos, seguida de un choque térmico a 345 K durante 10 minutos. Este proceso fue realizado gracias al uso de un termociclador (Mastercycler® de Eppendorf). Una vez inactivada la ligasa el ADN está listo para la transformación.

3.2.1.4. Transformación mediante choque térmico

Las ligaciones obtenidas en el punto anterior son transformadas en las células de clonaje XL1*blue* (tabla 3.1), las cuales son químicas-competentes. Se mezclan 10 µL de la reacción de ligación y 90 µL de las células competentes en esterilidad. Esta mezcla se incuba durante 30 minutos a 277 K, para a continuación aplicarle un choque térmico de 45 segundos a 315 K. Inmediatamente después, se incuban de nuevo a 277 K durante 4 minutos, para aplicarles 200 µL de SOC (*Super Optimal broth with Catabolite repression*, de Thermo Fisher Scientific). De esta forma las células se dejan a 310 K en agitación (800 rpm) durante una hora aproximadamente. Transcurrido este tiempo las células se plaquean en placas Petri con el medio Luria-Bertani (LB, de Melford) con agar al 1,5% más el antibiótico para la resistencia que porte el vector correspondiente. En la tabla 3.2 se reflejan los antibióticos usados, su concentración y el vector empleado con cada uno de ellos. Tras este proceso las placas se dejan a 310 K al menos 12 horas hasta la aparición de colonias.

Antibióticos	Concentración de trabajo (µg/mL)	Vector clonaje
Kanamicina	50	pETM11
Ampicilina	100	pET21b y pET23a

Tabla 3.2. Antibióticos empleados durante el proceso de clonaje y expresión de proteínas recombinantes y ARNt. Se muestran los antibióticos usados, kanamicina y ampicilina, así como las concentraciones empleadas durante los procesos de clonaje y expresión. Los vectores de expresión indicados son los que portan el gen de resistencia para cada uno de los antibióticos de la tabla.

Las colonias obtenidas fueron analizadas mediante una reacción de PCR. Para ello las colonias se resuspenden en 10 µL de agua estéril, de la cual se toman 5 µL como molde para la reacción de PCR. En este caso los oligonucleótidos empleados serán uno complementario al promotor portador del vector en el que se ha clonado, normalmente el promotor del bacteriófago T7 (extremo 5'), y el específico del gen clonado (extremo 3' del gen). La PCR es llevada del mismo modo que en el caso de amplificación del ADN del apartado 3.2.1.1. y analizado por electroforesis de ADN.

Una vez pasados estos controles, se secuencian las construcciones comprobadas por PCR, normalmente con los mismos oligonucleótidos mencionados en el caso de la PCR de colonia, con el fin de comprobar la secuencia obtenida nucleótido a nucleótido. Una vez confirmado que la construcción está bien y no ha sido introducida ninguna mutación al azar durante el proceso de clonación, se crecen dos de las colonias que dieron positivo en LB-agar y el antibiótico correspondiente y se aísla el ADN mediante una minipreparación (QIAGEN). El ADN obtenido de este proceso es usado para transformar la cepa de expresión BL21(DE3), del mismo modo al explicado para la cepa *XL1blue*.

3.2.1.5. Mutagénesis dirigida

La mutagénesis dirigida es una herramienta de biología molecular, la cual permite la manipulación de secuencias de ADN, ya sea para modificar un solo nucleótido, un codón o incluso eliminar zonas más extensas de este (deleciones). El fin último de esta técnica es el de manipular residuos o zonas específicas de una proteína con el fin de

entender mejor que relevancia tienen en la función de ésta. A continuación se explican los dos métodos por los cuales se realizó la mutagénesis dirigida de residuos específicos.

a) *Quik Change Site-Directed Mutagenesis*

Con el fin de desarrollar diferentes experimentos de la funcionalidad de algunas de las proteínas de esta tesis se tuvo que llevar a cabo la mutagénesis dirigida de residuos puntuales. Uno de los métodos empleados es el *Quik Change Site-Directed Mutagenesis II* (desarrollado por *Agilent Technologies*, www.genomics.agilent.com). Este método, basado en una reacción de PCR, simplemente necesita la construcción que codifica el gen a mutar, la cual se usa como molde de la reacción de PCR. Los oligonucleótidos empleados en este caso se complementan entre sí (tanto el directo como el reverso), como con el gen de la construcción, salvo por los nucleótidos que se quieran mutar, normalmente un codón, zona que debe estar en la zona intermedia de los oligonucleótidos. Estos oligonucleótidos deben tener una longitud de 40 bases aproximadamente y una temperatura de *melting* (T_m) de 350 K o más. En la figura 3.1 se muestra un esquema del proceso de mutagénesis mediante este método. En este tipo de reacciones la ADN polimerasa empleada fue la *Vent* de NEB.

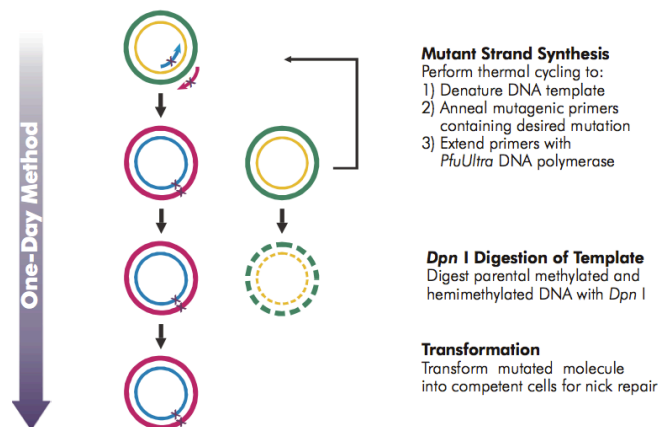


Figura 3.1. Representación esquemática del proceso de mutagénesis dirigida mediante *Quik Change Site-Directed Mutagenesis II*. Se representan los 3 pasos en este tipo de procedimiento, (i) PCR con el fin de introducir la mutación deseada, (ii) digestión con *DpnI*, eliminando el ADN molde metilado, y (iii) transformación en una cepa de clonaje de *XL1blue* (figura extraída de www.genomics.agilent.com).

Una vez finalizada la reacción de PCR, la muestra obtenida es analizada mediante electroforesis de ADN como se indicó anteriormente. Las muestras que dan amplificación son entonces incubadas con la enzima *DpnI* (NEB), la cual digiere solo el ADN metilado, durante una hora a 310 K. A continuación, y tras purificar el ADN obtenido (columnas QIAGEN), éste es transformado en la cepa *XL1blue* del modo explicado anteriormente, tras lo cual deben ser analizadas por secuenciación con el fin de confirmar la introducción de la mutación de forma correcta.

Otro de los aspectos a modificar son los ciclos empleados durante la reacción de PCR, diferentes a los convencionales, ya que hay que tener en cuenta que no se amplifica un inserto, si no la construcción completa. En primer lugar se aplica un ciclo de 370 K durante un minuto, para a continuación llevar a cabo de 12 a 16 ciclos de 30 segundos a 370 K, 1 minuto a 330 K y 1 minuto por kb de la construcción a 341 K. Finalmente se lleva a cabo un último paso a 341 K de 10 minutos.

b) Mutagénesis dirigida por empalme y extensión (*Site-directed mutagenesis by overlap extension*)

El otro método empleado con el fin de conseguir la sustitución dirigida de un solo residuo de una proteína fue el empalme y extensión (*site-directed mutagenesis by overlap extension*) mediante el uso de la técnica de PCR. En este método se hace necesario la producción de dos fragmentos de PCR (AB y CD) (figura 3.2) teniendo el gen diana como molde y empleando dos oligonucleótidos para cada fragmento, siendo uno de ellos el exterior del gen (cebador a o d) y el otro (el portador de la mutación) el interior de la secuencia del gen (cebador c o b). De esta forma, además se consigue que la secuencia original (sin la mutación) no se amplifique. No obstante tras la reacción de PCR, los fragmentos son digeridos con la enzima *DpnI* (NEB) con el fin de eliminar el ADN molde. Tras purificar los dos insertos generados, estos se emplean en la siguiente reacción de PCR como molde, y añadiendo en este caso solo los oligonucleótidos exteriores del gen (cebadores a y d). Como los dos insertos generados tienen una secuencia complementaria, en el primer ciclo de la PCR se obtiene el inserto completo con la mutación introducida, siendo los siguientes pasos de la reacción similares a cualquier otro caso de amplificación mediante PCR. Tras este paso simplemente se digiere el inserto con las enzimas de restricción para las que fueron diseñados los oligonucleótidos y se liga como se indica en el apartado anterior en el vector correspondiente.

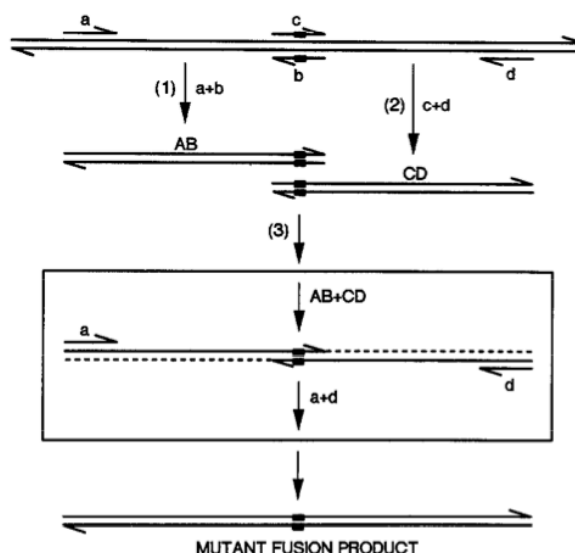


Figura 3.2. Diagrama esquemático del método de mutagénesis dirigida mediante empalme y extensión. El ADN de doble hebra así como los oligonucleótidos se representan mediante líneas, indicando con flechas la orientación 5' → 3'. Los sitios de mutagénesis se indican mediante rectángulos negros. Las letras minúsculas marcan los oligonucleótidos, mientras que las mayúsculas denotan los fragmentos de PCR. La caja representa el paso intermedio que tiene lugar en la reacción, donde los fragmentos intermedios, una vez desnaturizados, se empalman (*annealing*) para dar lugar al fragmento completo con la mutación introducida, mediante el uso de los oligonucleótidos exteriores (figura extraída de Ho *et al.*, 1989).

En este tipo de PCR la polimerasa empleada fue de nuevo la *Vent* (NEB) y el protocolo de amplificación fue similar al realizado en el apartado 3.2.1.1. En este caso tras finalizar el proceso es necesario llevar a cabo la ligación, tal y como se ha explicado anteriormente (apartado 3.2.1.3).

3.2.1.6. Pruebas de expresión a pequeña escala y posterior escalado

Una vez las construcciones han sido comprobadas y transformadas en la cepa de expresión BL21(DE3), se llevan a cabo las pruebas de expresión a pequeña escala, en 1 mL de medio de cultivo LB. Para ello, se crece biomasa del clon en la cepa de expresión con el antibiótico para el que es resistente a 310 K hasta alcanzar una densidad óptica a 590 nm ($D.O_{590}$) de 0,6-0,8. A continuación se induce con 0,5 mM de IPTG (Isopropil- β -D-1-tiogalactopiranosido) durante 5 horas a la misma temperatura. Una vez finalizado la expresión, el cultivo se centrifuga a 2.000 g (277 K 5 minutos), decantando el sobrenadante y resuspendiendo el sedimento celular generado en 400 μ L del tampón

PBS pH 7,0 (Tampón Fosfato Salino: 140 mM NaCl, 3 mM KCl, 10 mM Na₂HPO₄ y 2 mM KH₂PO₄). Esta muestra es sonicada durante 30 segundos a 277 K al 70% de potencia. Una vez lisado el cultivo se toman 15 µL de la muestra y se centrifuga el resto a 10.000 g (277 K 10 minutos), para tomar a continuación 15 µL del sobrenadante. Estas muestras son mezcladas con 5 µL de tampón de carga (100 mM Tris, 2 mM EDTA, 8% SDS, 40% glicerol, 4 mM β-mercaptoetanol y 0,8% *bromophenol blue*) cada una y calentadas a 345 K durante 10 minutos. Las muestras son analizadas mediante PAGE-SDS, empleando además 5 µL del marcador de proteínas (Pierce™ Unstained Protein MW Marker). Los geles se tiñeron con Azul Brillante de Coomassie (CBB, de *Coomassie Brilliant Blue*). En el apartado 3.2.1.4. se detalla la preparación de geles para PAGE-SDS y del procedimiento para llevar a cabo la electroforesis.

Una vez comprobada la expresión de la proteína recombinante a pequeña escala, se procede al escalado de la misma si la cantidad y la calidad son aceptables, o se procede a la optimización de la expresión si se observa poca cantidad o degradación. Los parámetros a optimizar son: (i) la cantidad de inductor empleado, (ii) el tiempo de expresión o (iii) la temperatura de expresión, principalmente, aunque existen otros procedimientos como pueden ser clonajes en otros vectores o cambios a medios de cultivo más ricos en el proceso de expresión.

Una vez encontradas las condiciones óptimas, se procede al escalado de la expresión, trabajando con 1 litro de cultivo o más, en matraces de 2 litros de capacidad, donde no se añade más de la mitad de capacidad del recipiente con medio de cultivo. Con el fin de iniciar la expresión primero se necesita crecer biomasa hasta alcanzar una D.O.₅₉₀ de 0,6-0,8 a 310 K (800 rpm) en el cultivo. Una vez alcanzada la D.O.₅₉₀ indicada, se induce durante el tiempo óptimo con la concentración de IPTG elegida. Finalizado el tiempo de expresión, el cultivo celular es centrifugado a 2.000 g (277 K 15 minutos). Posteriormente se decanta el sobrenadante para resuspender el sedimento celular en PBS a pH 7,0. A continuación se vuelve a centrifugar en las mismas condiciones para decantar de nuevo el sobrenadante y guardar el sedimento celular a 195 K hasta su procesado.

3.2.1.7. Producción y obtención ARN de transferencia

En el caso de la expresión de los ARNt clonados, se llevó a cabo un protocolo similar al anterior, variando el volumen de medio empleado, siendo en este caso de 200 mL de LB más el antibiótico correspondiente. El proceso de inducción fue similar (empleando IPTG como inductor), mientras que el proceso de lisis celular difiere al anterior. Una vez obtenido el sedimento celular con la sobreexpresión del ARNt, éste es resuspendido en el tampón de lisis, el cual ha sido previamente autoclavado para

minimizar la acción de ARNasas. En la medida de lo posible, las muestras con ARNt siempre se conservan a 277 K. Posteriormente se añade un mismo volumen de fenol (pH 4,3) al sedimento celular resuspendido y se agita y centrifuga la muestra a 10.000 g (10 minutos 277 K). Tras este proceso se toma la fase acuosa y se mezcla con 2,5 volúmenes de etanol al 100%, agitándose e incubándose (1 hora 277 K). Posteriormente la muestra es centrifugada a 20.000 g (20 minutos 277 K). El sedimento generado es resuspendido en el tampón de la filtración en gel (técnica explicada más adelante). Las muestras obtenidas son analizadas en este caso mediante un gel de agarosa al 2%, previamente mezclado con SYBR® *Safe*, como se indica en el apartado 3.2.1.2. En este caso, debido al tamaño de la molécula de ARNt el marcador empleado es NZYDNA *Ladder* VI.

3.2.1.8. Electroforesis de proteínas en condiciones desnaturizantes (PAGE-SDS, *Polyacrylamide Gel Electrophoresis-Sodium Dodecyl Sulfate*) y nativas

Una vez han sido preparadas las muestras de proteína, ya procedan de las pruebas de expresión o de cualquier proceso de purificación, éstas son cargadas en un gel de acrilamida para llevar a cabo su análisis en condiciones desnaturizantes (PAGE-SDS).

Los geles empleados poseen un porcentaje de acrilamida del 10% en la zona del *stacking* (zona superior del gel donde se cargan las muestras) y del 15% en la zona de *resolving* (zona de resolución). El gel se corre en una fuente con tampón (25 mM Tris, 200 mM Glicina y 0,1% (v/v) SDS), a 250 V durante 30 minutos aproximadamente a temperatura ambiente. El gel es teñido con azul de *Coomassie*, acelerando el proceso de tinción mediante sucesivos calentamientos en un microondas.

En cuanto a los geles en condiciones nativas de acrilamida, éstos fueron preparados al 8% de acrilamida, empleando para ello en este caso el tampón TBE (0,5X), con el fin de estudiar complejos proteína-ARN en condiciones nativas. La carrera fue llevada a cabo en una cubeta para tal fin a 12 mA, 277 K, durante 1 hora. El gel fue teñido con primeramente con *Coomassie Brilliant Blue* (con el fin de visualizar la proteína) y después con *Toluidine Blue* (con el fin de visualizar el ARN).

3.2.2. Purificación de proteínas

3.2.2.1. Lisis celular

A la hora de procesar los sedimentos procedentes de la expresión celular, estos fueron almacenados a 195 K hasta su procesamiento. Para ello, eran descongelados a 277 K y resuspendidos en el tampón de lisis, el cual lleva añadido 1 mM de PMSF

(fluoruro de fenilmetilsulfonilo). Esta muestra, una vez homogenizada es sonicada al 70% de potencia durante 2,5 minutos en ciclos de 30 segundos de sonicación y 20 segundos de reposo. Este método de ruptura de células se da gracias al proceso de cavitación y de *liquid shear disruption*, permitiendo la liberación del contenido celular. A continuación la muestra es centrifugada a 10.000 g (20 minutos 277 K). Posteriormente el sobrenadante generado es filtrado mediante el empleo de filtros de 0,45 µm de PES (Polietersulfona) (Grabski, 2009).

3.2.2.2. Medición de concentración de proteína y ácido nucleicos

Con el fin de cuantificar la concentración de proteínas o ácido nucleico que se tenía se realizaron mediciones de las muestras mediante espectrofotometría, a 280 nm para el caso de las proteínas, y a 260 nm para el caso de los ácidos nucleicos. Para ello se empleó el espectrofotómetro *Eppendorf BioSpectrometer* y las cubetas de cuarzo (*Micro Quartz cuvette*). Del mismo modo se realizaba el blanco en cada caso con el tampón en el cual se encuentra la muestra a medir. En el caso de las proteínas se empleaba además el coeficiente de extinción teórico obtenido mediante el portal ExPaSy.

3.2.2.3. Purificación por cola de afinidad

La mayoría de las proteínas recombinantes sobreexpresadas, son modificadas mediante ingeniería genética con el fin de facilitar el proceso de purificación y/o detección. Para ello normalmente se les añade una cola de afinidad o bien en el extremo N-terminal o en el C-terminal.

Una de las colas de afinidad más aplicadas hasta el momento es la cola de histidinas, compuesta de 6 o más residuos de este aminoácido. Debido al pequeño tamaño de esta cola y al no estar cargada a un pH fisiológico, permite mantener la estabilidad del plegamiento de la proteína recombinante, además de no interferir, en la mayoría de los procesos, en su función.

La purificación por cola de histidinas es comúnmente conocida como IMAC (*Immobilized Metal Ion Affinity Chromatography*), pues en la matriz que compone la columna de purificación empleada se encuentra un quelante capaz de unir iones de metal. Esta matriz normalmente está compuesta de NTA (ácido de nitrilotriacético), la cual es capaz de unir tanto iones de níquel, el metal más usado en este tipo de purificaciones, como de cobre. De esta forma las proteínas modificadas para portar una cola de histidinas son capaces de unirse de forma específica a este tipo de resinas modificadas a través del metal unido.

Las purificaciones de tipo IMAC son procesos de unión a la columna, lavado, con el fin de eliminar impurezas que se hayan unido de forma inespecífica, y elución de la proteína recombinante. Este último paso consiste en una competición entre la proteína unida y el imidazol que se aplica en el tampón de elución, el cual hará que se libere la proteína unida específicamente. Además debido al alto grado de afinidad de este tipo de procedimiento y a la especificidad de la cola de histidinas, un solo paso de purificación suele ser suficiente, dependiendo del grado de pureza que se exija para la muestra (Block *et al.*, 2009).

En cuanto a los equipos empleados para ello, se suele emplear equipos de HPLC (*High-Performance Liquid Chromatography*), el cual permite añadir diferentes tampones, por ejemplo el de lavado o elución en el caso de la IMAC, así como aplicar gradientes de forma precisa, controlando parámetros como el flujo o la presión de la muestra que atraviesa la columna. Alternativamente, la purificación por IMAC se puede realizar por gravedad o con la ayuda de una bomba peristáltica (*P1 pump*, GE Healthcare).

En todos estos casos, tras finalizar la elución de las proteínas, las muestras son analizadas mediante PAGE-SDS antes de continuar con el tratamiento de la proteína.

3.2.2.4. Intercambio iónico

Este método de purificación es el más popular, pues puede ser aplicado a cualquier proteína, ya que explota una propiedad que toda molécula proteica posee, motivos de cargas superficiales. Éstos son capaces de interaccionar con intercambiadores de iones, lo que permite su unión a la columna y posterior elución. Además este paso de purificación, como los anteriores, permite concentrar la proteína diana en la columna de unión, pues es retenida toda de forma específica en un volumen menor.

En cuanto al proceso de purificación, la proteína se une a la columna en función de su carga neta, lo cual dependerá del pH de la solución en la que se encuentre y de su punto isoeléctrico. De esta forma, si nos encontramos a un pH por debajo del punto isoeléctrico de la proteína, ésta presentará una carga neta negativa, mientras que si el pH se encuentra por encima tendrá carga neta positiva. En cuanto al proceso de elución, éste se realiza aumentando la concentración de sal del tampón, de forma que competirá con la proteína por unirse a la columna. Este proceso se suele hacer en gradiente para obtener proteína de mayor pureza, gracias al empleo de equipos de tipo HPLC.

En cuanto al intercambiador de iones usado, éste se compone de una matriz, a la cual se une un ligando cargado, ya sea positiva o negativamente. Así los intercambiadores de cationes se encuentran cargados negativamente, por lo que

intercambian proteínas cargadas positivamente en un proceso denominado intercambio catiónico. Por el contrario, los intercambiadores de aniones se encuentran cargados positivamente, por lo que unirá proteínas con carga neta negativa, en un proceso denominado intercambio aniónico (Jungbauer y Hahn, 2009).

De nuevo, este tipo de purificaciones son del tipo unión, lavado y elución, donde la unión y el lavado se realizan a una concentración baja de sal, mientras que en la elución se aplica un gradiente de sal, aumentando gradualmente la concentración de ésta hasta conseguir la elución de la proteínas diana.

En todos estos casos la calidad de la proteína purificada es analizada mediante PAGE-SDS.

3.2.2.5. Filtración en gel

Este tipo de purificación, también conocido como SEC (*Size Exclusion Chromatography*), se basa en la separación de los componentes de una solución en función de su tamaño, en concreto de su radio hidrodinámico. De esta forma, a la vez que se lleva a cabo la purificación de proteínas, también es posible la determinación del estado oligomérico de complejos proteicos, así como intercambiar el tampón en el que se encuentra disuelta una proteína, actuando en este último caso como columna de desalado.

Las columnas en las cuales se lleva a cabo este tipo de purificación, empleando normalmente equipos de HPLC, están formadas por una matriz de esferas que poseen un tamiz de un tamaño de poro determinado para cierto tamaño molecular. Esta matriz puede estar formada de poliacrilamida, agarosa, dextrano o alguna combinación de los materiales anteriores.

Las columnas destinadas a este tipo de purificación se caracterizan por el tamaño de los poros del tamiz, el volumen muerto y el volumen de la columna. El primero de ellos, el tamaño del poro, equivale al tamaño aproximado de la molécula más grande que se puede introducir en las esferas que componen la columna. Cualquier molécula o especie de mayor tamaño quedará en el volumen muerto y no será capaz de separarse por su tamaño. De esta forma el volumen muerto se define como el volumen que no está ocupando la matriz dentro de la columna. Por su parte, el volumen de la columna se refiere al volumen total accesible dentro de ésta. Así, moléculas de pequeño tamaño como las sales entrarán en la matriz y eluirán aproximadamente con el volumen de la columna. Por lo tanto, las moléculas de mayor tamaño saldrán antes mientras que las de

menor tamaño saldrán más tarde, de acuerdo con su tamaño molecular, ya que son retenidas dentro del tamiz.

En este caso las moléculas de proteínas en ningún momento se unen a la fase estacionaria, hecho que la distingue de las anteriores técnicas de purificación, donde sí hay unión. Del mismo modo, la filtración en gel suele ser un paso de purificación complementario que se realiza en último lugar, pues no es aconsejable para un lisado celular, ya que existen múltiples proteínas de parecido peso molecular, por lo que no sería posible distinguirlas de forma adecuada (Stellwagen, 2009).

3.3. Métodos de análisis

3.3.1. *Cross-linking*

Cross-linking o entrecruzamiento es el proceso por el cual se produce una unión química entre dos o más moléculas de forma covalente. Los *crosslinkers* (entrecruzadores) son moléculas que contiene extremos reactivos capaces de unirse de manera específica a determinados grupos químicos, como pueden ser aminas primarias o grupos sulfhidrilo (-SH), ya sea en proteínas u otro tipo de moléculas. Este tipo de experimentos se pueden realizar con distintos fines. Uno de ellos es la estabilización de la unión de proteínas, ya sea en su estructura terciaria o cuaternaria, mediante la formación de puentes intramoleculares o intermoleculares (entre las subunidades de un complejo proteico). Otra de las aplicaciones es la captura e identificación de interactores (proteínas que interaccionan) desconocidos de una proteína dada, mediante el posterior análisis. Además esta técnica también puede ser empleada para la conjugación de diferentes moléculas a enzimas o a anticuerpos u otro tipo de moléculas con el fin de marcarlas. Del mismo modo se puede lograr la inmovilización de anticuerpos u otro tipo de proteínas a resinas con el fin de llevar a cabo procesos de purificación por afinidad.

En nuestro caso usamos el *crosslinker* BMOE (Bis(Maleimido) Etano) (Figura 3.3), que es capaz de conjugar grupos sulfhidrilo, en concreto dos por molécula de BMOE. La reacción entre el BMOE y el grupo sulfhidrilo da como resultado un enlace tioéter estable, el cual no puede ser reducido por agentes reductores o roto bajo condiciones fisiológicas de pH, además de caracterizarse por ser una reacción de alta especificidad. Esta molécula de BMOE se caracteriza por tener un peso molecular de 220,18 Da, dato importante a la hora de los estudios de interacción con otras moléculas en cuanto a la identificación mediante las masas generadas. Además el brazo espaciador de la

molécula es de 8,0 Å, dato importante pues refleja la distancia a la que se pueden encontrar en la naturaleza los residuos sulfhidrilo que intervienen en la interacción (www.thermofisher.com).

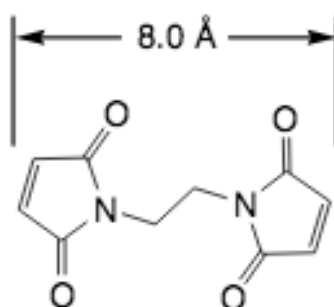


Figura 3.3. Molécula del crosslinker BMOE. Se muestra la estructura molecular del BMOE, así como la distancia (8,0 Å) a la que quedan las moléculas que une. Imagen extraída de www.thermofisher.com.

3.3.2. DLS (*Dynamic Light Scattering*)

DLS es una técnica que se usa de forma rutinaria en los laboratorios de biología estructural con el fin de determinar el tamaño de moléculas de proteína o de complejos, así como de ácidos nucleicos, además de monitorizar la unión de ligandos. Del mismo modo es posible detectar la aparición de diferentes agregados en soluciones de proteína.

Esta técnica permite, en condiciones nativas, determinar el coeficiente de difusión translacional (D_t), también denominado difusividad de los componentes de una solución. D_t describe la facilidad que posee una sustancia para desplazarse en el interior de otra mediante el fenómeno de difusión. El D_t puede reflejar el tamaño de las partículas pues es inversamente proporcional al tamaño de éstas, moviéndose las partículas de menor tamaño con mayor rapidez que las de mayor tamaño (Lorber *et al.*, 2012).

En cuanto a las soluciones compuestas por una sola población de partículas, por ejemplo la muestra obtenida de un último paso de purificación de una proteína determinada, el D_t nos puede dar dos tipos de información. En primer lugar, el tamaño de partícula, en cuanto al diámetro hidrodinámico, asumiendo que las partículas de la

solución son esferas. Y segundo, la polidispersidad de la población en solución, obtenida de la desviación estándar del D_t , dato que muestra la calidad de la muestra a la hora de realizar un experimento de cristalización, siendo aceptables valores de polidispersidad por debajo del 30%, lo que indica el carácter homogéneo de la muestra en cuanto a su composición (Shiba *et al.*, 2010).

Por último cabe destacar la extrema sensibilidad de esta técnica frente a la detección de agregados, lo cual es de gran utilidad en biología estructural, pues permite la identificación de muestras de proteína en las cuales las macromoléculas muestran monodispersidad, esencial para la formación de núcleos de cristalización y el crecimiento de los cristales de proteína.

3.3.3. SPR (*Surface Plasmon Resonance*) o resonancia de plasmón de superficie

Con el fin de estudiar la interacción entre proteínas expresadas se llevaron a cabo ensayos de unión mediante la técnica SPR. Esta técnica permite estudiar la interacción entre proteínas y otros ligandos. Además se caracteriza por ser libre de marcaje de ninguno de los componentes, además de necesitar de pequeñas cantidades de muestra para el ensayo. Para llevar a cabo un experimento de SPR es necesario inmovilizar uno de los componentes (el ligando) a la superficie de dextrano de un chip CM5 (BIAcore, GE Healthcare) en nuestro caso. El chip posee una fina capa de oro sobre el lado opuesto a la unión de los componentes a ensayar. La inmovilización puede ser covalente o indirecta, mediante la inmovilización covalente de un anticuerpo, en el chip, capaz de capturar el ligando. En nuestro caso, se inmovilizó un anticuerpo monoclonal capaz de reconocer específicamente colas de histidinas, a la cual se unirá una de las proteínas a estudiar. Una vez inmovilizados los componentes al chip se deja fluir por la superficie de éste una solución con el otro componente a estudiar (el analito). Al mismo tiempo la superficie del chip que posee la capa de oro es irradiada con un haz de luz, explotando de esta forma el fenómeno de generación de plasmones, de tal forma que dependiendo de la cantidad de analito unido al ligando, el grosor del chip cambiará dando lugar a cambios en el índice de refracción de la superficie, y por lo tanto el ángulo de incidencia del haz de luz (θ) también se verá afectado en proporción, pudiendo medir de esta forma constantes de afinidad de los componentes (Patching *et al.*, 2014). En la figura 3.4 se representan los principios de esta técnica.

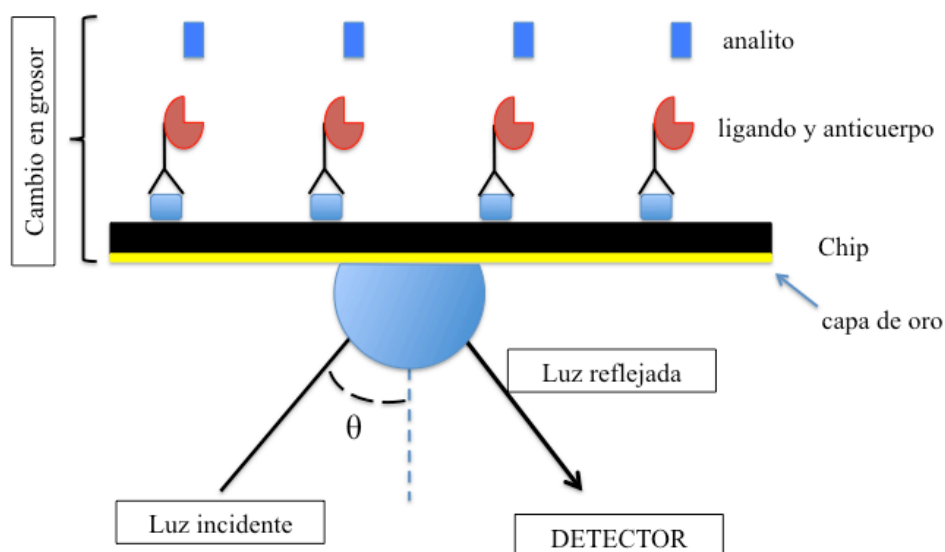


Figura 3.4. Esquema representativo de los principios de SPR. En este caso se muestra la unión covalente de un anticuerpo, capaz de reconocer un ligando de forma específica, unido covalentemente al chip. La unión del analito a estudiar influirá en el grosor del chip y por lo tanto, en el ángulo de incidencia (θ), permitiendo calcular, gracias al detector, constantes de afinidad.

En nuestro caso se inmovilizó la proteína con cola de histidinas sobre un chip CM5. Éste primeramente se trató con un anticuerpo anti-cola de histidinas (α -His kit, GE Healthcare), quedando el anticuerpo fijado al chip mediante una amina acoplada químicamente (Amino Coupling kit, GE Healthcare). El instrumento empleado para llevar a cabo los experimentos fue BIAcore X100 (GE Healthcare).

3.3.4. Espectrometría de masas

La espectrometría de masas es un método empleado para la identificación de proteínas o péptidos. La identificación de proteínas se basa en la generación de péptidos a partir de la digestión con proteasas de éstas. Para ello, la enzima más usada es la tripsina, la cual es capaz de cortar la cadena peptídica en el lado C-terminal de residuos con carga positiva, como es el caso de lisina o arginina, salvo que el siguiente aminoácido en la secuencia sea una prolina. Este podría ser el caso de una muestra conocida. Sin embargo, cuando se posee una muestra heterogénea en cuanto a la composición de proteína se analizan las muestras mediante PAGE-SDS con el fin de separar cada proteína, la cuales se pueden obtener posteriormente del gel de

acrilamida. Seguidamente estas muestras pueden ser procesadas, tras su tripsinización, mediante MALDI-TOF (*Matrix-Assisted Laser Desorption/Ionization, Time Of Flight*). En esta técnica se produce la ionización de las proteínas o péptidos obtenidos, de tal forma que las molécula a analizar pasan del estado sólido al gaseoso. A continuación los iones generados son acelerados midiendo su tiempo de vuelo, obteniendo resultados de las muestras que darán valores de la masa y carga (m/z) de los distintos péptidos, siendo la masa proporcional al valor obtenido y la carga inversamente proporcional. Finalmente los datos obtenidos de los valores de m/z son analizados con diferentes programas y comparados con bases de datos con el fin de identificar las diferentes proteínas (Finehout y Lee, 2004).

3.3.5. Actividad enzimática de OPRT

Con el fin de estudiar las constantes cinéticas de algunas de las proteínas de esta tesis se llevaron a cabo el estudio de su actividad enzimática. Para ello se midió la absorbancia de uno de los sustratos de la reacción a una longitud de onda de 295 nm, frente al tiempo, fijando el resto de los componentes de la reacción (otros sustratos así como la concentración de la enzima). La temperatura a la cual fueron llevados los experimentos también fue fijada a 298 K. El resultado de estas mediciones dio datos de absorbancia/tiempo, datos transformados por la aplicación de la ecuación de Lambert-Beer ($A = \varepsilon \cdot c \cdot l$, donde **A** representa la absorbancia a una longitud de onda dada, ε el coeficiente de extinción molar, **c** la concentración y **l** el paso óptico). De esta forma se originaron datos de concentración/tiempo, con los cuales se pudieron calcular las constantes cinéticas de K_M , V_{\max} y k_{cat} gracias al uso del programa SigmaPlot 13.0.

3.4. Métodos de caracterización estructural

3.4.1. AUC, ultracentrífuga analítica

Este método es una potente herramienta para el análisis cuantitativo de las propiedades hidrodinámicas de macromoléculas en solución, como su tamaño, masa y forma, así como para el estudio de las interacciones entre macromoléculas. La AUC se basa en el análisis del desplazamiento de las macromoléculas biológicas a lo largo de una columna de líquido (su sedimentación) bajo la acción de un campo de aceleración generado durante la centrifugación de las muestras. El desplazamiento de las moléculas se monitoriza mediante tres posibles sistemas (absorbancia, fluorescencia o interferencia de masas). En esta tesis se ha empleado absorbancia a 280 ó 260 nm.

Existen dos procedimientos experimentales principales en AUC. El primero es velocidad de sedimentación (SV), que proporciona información hidrodinámica sobre el tamaño y la forma de las macromoléculas (Howlett *et al.*, 2006; Laue y Stafford, 1999; Lebowitz *et al.*, 2002). La segunda aproximación, equilibrio de sedimentación (SE), es un método termodinámico que emplea el análisis de los gradientes de concentración en el equilibrio para derivar información sobre las masas molares en solución, estequiometría, constantes de asociación/disociación (Howlett *et al.*, 2006; Laue, 1995). Ambos tipos de experimentos requieren protocolos especializados que se implementan utilizando instrumentos conocidos como ultracentrífugas analíticas. En ambos casos, al someter la muestra al campo centrífugo, se produce una redistribución de la masa a lo largo de la celda de muestra hasta que la energía potencial debido a la aceleración centrífuga equilibra exactamente el potencial químico en cada posición radial. Si monitorizamos la velocidad a la cual se mueven las bandas de moléculas durante esta redistribución, estamos realizando un experimento de SV. Si, en cambio, determinamos la concentración en cada posición radial en el equilibrio, estamos realizando un experimento de SE (Cole *et al.*, 2008).

En los experimentos de SV se obtiene información de la masa molecular de las partículas a analizar, además de su forma. En este tipo de experimento la muestra es centrifugada a una gran velocidad, del orden de 40-60 krpm, y se toman medidas de la distribución radial de concentración del soluto a intervalos cortos (minutos). Cada medida se denomina un “scan”.

En los experimentos de SE la velocidad de centrifugación es mucho menor, para generar un gradiente de concentración a lo largo de la celda de medida estable en el tiempo como resultado del balance exacto entre la sedimentación y la difusión (Cole *et al.*, 2008; Lebowitz *et al.*, 2002).

3.4.2. Resonancia magnética nuclear (RMN)

La resonancia magnética nuclear es un fenómeno que ocurre cuando el núcleo de cierto átomo es afectado por un campo magnético, generado por un imán de gran potencia, y es expuesto a una oscilación por parte de éste.

Para efectuar el experimento con proteínas, estas son marcadas con isótopos como el ^{13}C o el ^{15}N , los cuales pueden ser fácilmente incorporados en el proceso de expresión de las mismas, al aplicar D-glucosa con ^{13}C y cloruro de amonio con ^{15}N en medios de cultivo mínimos durante el proceso de producción de la proteína recombinante.

La frecuencia exacta de resonancia depende del ambiente químico del átomo, de tal forma que para cada experimento el espectro de RMN de una proteína dada mostrará las señales de resonancia con leves diferencias de frecuencia, lo que se conoce como desplazamiento químico. Exactamente lo que se mide en RMN son las distancias, por ejemplo, protón-protón, así como los ángulos, lo cual pueden llevar, tras un proceso computacional, a determinar estructuras de proteínas hasta 15 kDa (Gao *et al.*, 2004). La RMN aplicada a la biología puede aportar información sobre la dinámica conformacional de biomoléculas, además de ser muy eficiente en la determinación de la unión de ligandos, así como del mapeo de las superficies de interacción de proteínas con sus ligandos o interactores. El primer paso en la determinación estructural mediante el uso de RMN consiste en la asignación de los desplazamientos químicos de todos los átomos de la molécula de estudio, gracias al espectro de RMN generado. Una vez estos átomos son identificados, se procede a definir la estructura tridimensional de la molécula. Es así como se pueden identificar que zonas de una proteína intervienen en la interacción con otras moléculas, al observar esos desplazamientos químicos en ciertos residuos y no en otros (Gao *et al.*, 2004). Para estos estudios existen diferentes experimentos. Los empleados en esta tesis fueron de dos tipos. El primero de ellos, donde una de las proteínas se encontraba doblemente marcada (^{13}C y ^{15}N) fue de tipo 1D ^1H - ^{13}C HSQC (*Heteronuclear Single-Quantum Correlation spectroscopy*) y ^1H - ^{15}N HSQC. Estos experimentos son monodimensionales, por lo cual se midió la resonancia procedente o de los núcleos de nitrógeno o de carbono de la proteína marcada, al interaccionar con otra molécula no marcada. En este caso no es posible asignar la señal a los diferentes átomos de la proteína. Sin embargo sí permite ver diferencias en los espectros generados entre la proteína marcada sola o interaccionando con la no marcada.

Por otro lado, los experimentos de tipo NOESY (*Nuclear Overhauser Effect Spectroscopy*), donde de nuevo una de las proteínas estaba doblemente marcada, son bidimensionales. En este caso se explota la interacción dipolar entre los *spins* de ^1H y ^{15}N de la proteína marcada, además de ayudarse en este caso de la estructura resuelta mediante RMN de la proteína marcada. Aquí sí es posible asignar cada señal observada en el espectro con una amida de la proteína, por lo cual es posible detectar cambios en la proteína a estudiar, al interaccionar con otra proteína no marcada, en residuos concretos. En la figura 3.5 se muestra un espectro bidimensional donde es posible asignar la señal a residuos específicos de la proteína marcada.

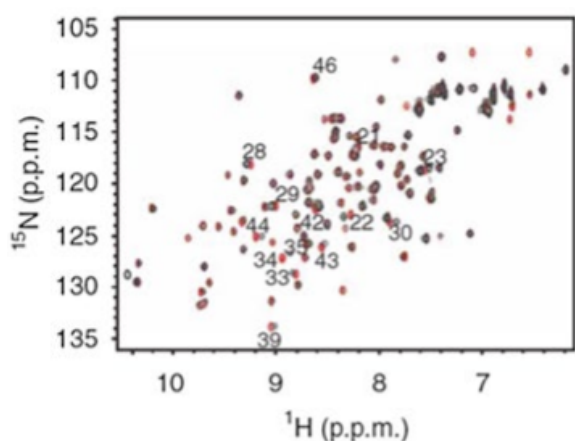


Figura 3.5. Experimento de NMR de tipo NOESY donde se comparan los espectros generados por una proteína marcada con ^{15}N al incubarla con otra proteína sin marcar. Se marcan los residuos donde se encuentran diferencias en cuanto a su resonancia. En el eje Y se muestran las p.p.m. (partes por millón) procedentes del ^{15}N , mientras que en el eje X se indican las p.p.m. procedentes del ^1H . Imagen extraída de Prischi y Pastore, 2016.

Los experimentos de RMN se realizaron en colaboración con el grupo del Dr. Jiménez Barbero (Departamento de Biología Físico-Química, Centro de Investigaciones Biológicas (CSIC) (Madrid, España) y Departamento de Biología Estructural, CIC bioGUNE Derio-Bizkaia, España).

3.4.3. SAXS (*Small Angle X-ray Scattering*)

El método SAXS, o difracción de rayos-X de bajo ángulo, permite estudiar la estructura y propiedades hidrodinámicas de proteínas y macromoléculas en disolución mediante el uso de la difracción de rayos-X. A diferencia de la cristalografía de rayos-X, las muestras consisten en disoluciones de macromoléculas y, por tanto, no es necesario cristalizar la muestra. Como las macromoléculas en disolución se encuentran dispuestas en todas las orientaciones y posiciones posibles, la imagen de difracción captada en el detector de rayos-X se corresponde con un promedio rotacional y traslacional de todas las moléculas en disolución (Putman *et al.*, 2007). Cuando se irradia una muestra líquida con rayos-X se obtienen patrones de difracción que son simétricos radialmente y que pueden integrarse para dar lugar a una curva monodimensional de intensidad de difracción frente a $1/(\text{resolución})$.

A partir de los datos de SAXS se pueden derivar varios parámetros importantes de la estructura global de las macromoléculas y que no necesitan asumir ningún tipo de restricción sobre las propiedades macromoleculares. La primera de ellas, la ley de Guinier, permite calcular el radio de giro (R_g) de la macromoléculas en disolución. A partir del R_g y aplicando suposiciones simplificadoras se podría estimar el tamaño total

de la molécula de estudio. Aplicando la misma ley, se puede determinar la intensidad de la dispersión a ángulo cero [$I(0)$], que es proporcional a la masa molecular, y comparando con un estándar de masa molecular conocida (típicamente, albúmina de suero bovino, BSA) se puede extrapolar una masa molecular total para la especie en disolución. Un segundo parámetro importante es la dimensión máxima (D_{\max}), que permite caracterizar cuál es la máxima distancia intramolecular y, por tanto, informa sobre la forma molecular (*e.g.*, esférica, alargada, discoidal). Otras transformaciones de los datos permiten calcular el volumen de hidratación de la macromolécula, a través de la ley de Porod, y de ahí estimar la masa molecular, así como el estado de plegamiento de la macromolécula a través de la representación de Kratky (Round, 2016). En la figura 3.6 se muestra un ejemplo de los modelos generados a partir de experimentos de SAXS.

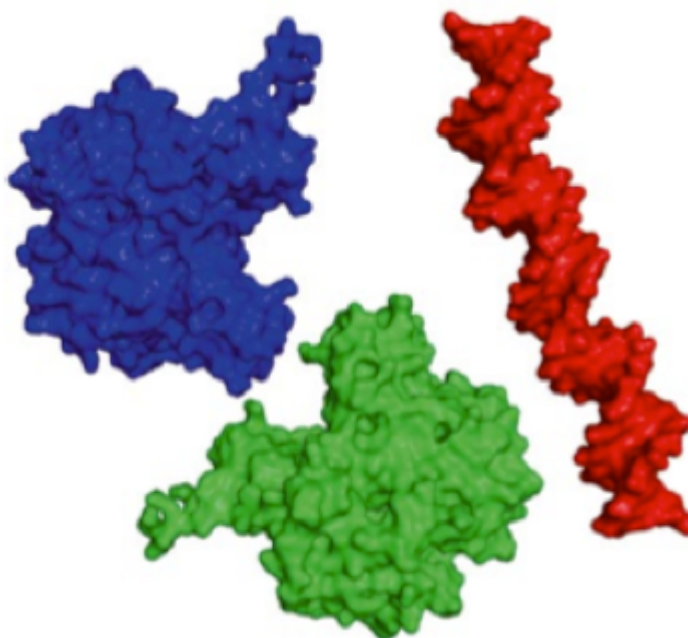


Figura 3.6. Modelo generado por SAXS de tres proteínas a partir de los datos obtenidos experimentalmente. Se muestra la reconstrucción de las máscaras de tres proteínas diferentes a partir de SAXS, lo que permite el estudio de sus estructuras tridimensionales (imagen extraída de Round, 2016).

Los experimentos de SAXS se realizaron en colaboración con el Dr. Round (Laboratorio Europeo de Biología Molecular, Grenoble, Francia y Unidad para la

interacción de virus células huésped, Universidad de Grenoble Alpes-EMBL-CNRS, Martyrs, Francia).

3.5. Métodos de cristalografía de rayos-X

3.5.1. Métodos de cristalización

El principal requerimiento para resolver la estructura tridimensional de una proteína mediante cristalografía de rayos-X es la obtención de cristales de proteína de calidad. El proceso de cristalización de macromoléculas no se puede predecir ya que el mecanismo de nucleación es de naturaleza estocástica. Esto ocasiona que sea necesario ensayar multitud de condiciones de cristalización hasta optimizar el proceso. Un dato importante a la hora de obtener cristales de calidad es que la proteína precipite lentamente y de manera ordenada.

Hay cinco pasos necesarios para la cristalización de proteínas:

1. Obtener la proteína en el mayor grado de pureza posible. A la hora de lograr que las proteínas cristalicen es esencial que todas las moléculas tengan las mismas propiedades superficiales y que la solución sea lo más homogénea posible.
2. La proteína debe estar solubilizada, durante el experimento de cristalización, en un tampón adecuado.
3. Las condiciones de cristalización más frecuentemente empleadas con éxito suelen contener combinaciones de sal (por ejemplo, sulfato de amonio) y/o PEG (polietilenglicol), o compuestos orgánicos como MPD (2-metil-2,4-pentanodiol). Estos compuestos están entre las condiciones en que se han obtenido cristales en esta tesis.
4. La solución de proteína debe alcanzar un estado de supersaturación. Esto permite que se formen núcleos de cristalización (nucleación), pequeños agregados que serán la base para el crecimiento de cristales. A medida que se incorporan nuevas moléculas de proteína a los núcleos, la energía de tensión superficial disminuye, favoreciéndose el crecimiento de los núcleos.
5. Una vez formados esos núcleos se produce el crecimiento real de los cristales, hasta alcanzar un tamaño óptimo para ser difractados mediante rayos-X.

Existen varias formas para modular el grado de supersaturación y por lo tanto tener más posibilidades de obtener mejores cristales. Se pueden variar parámetros como la temperatura a la cual se busca que se formen los cristales, la concentración del agente precipitante o modificar el pH de la solución, además de variar la concentración de proteína empleada para la cristalización. La figura 3.7 se muestra la relación entre la concentración de proteína y de agente precipitante a la hora de formar núcleos para la formación de cristales (Drenth, 2007).

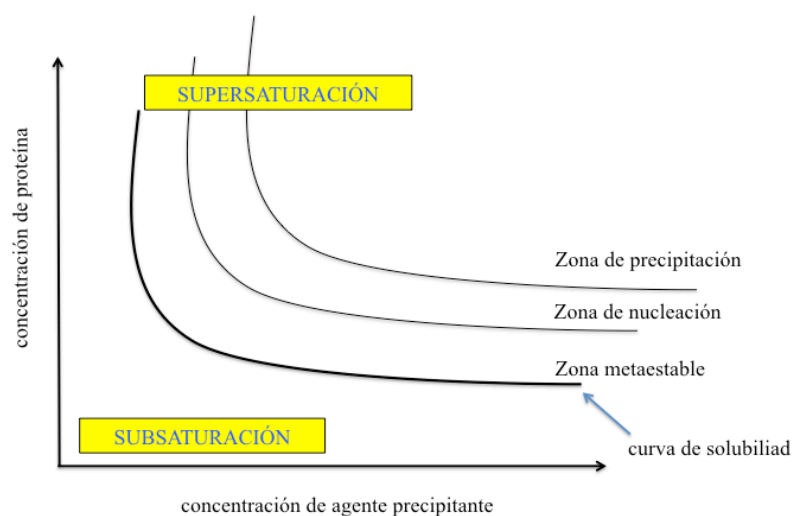


Figura 3.7. Diagrama de fases bidimensional. Se ilustra el cambio en la concentración de proteína frente al cambio en la concentración del agente precipitante. Se puede observar que la curva de solubilidad se encuentra entre las zonas de subsaturación y supersaturación de una solución proteica, ésta última abarcando las zonas metaestable, de nucleación y de precipitación.

En el esquema del diagrama de fases (figura 3.7), los cristales de proteína jamás crecerán en la zona de subsaturación. En cuanto a la zona de supersaturación, ésta se divide en tres zonas. La primera de ellas es la de precipitación, donde el exceso de moléculas de proteína genera agregados amorfos. El segundo es el conocido como zona de nucleación, donde el exceso de proteína tiende a la formación de núcleos, base para la formación de los cristales. Por último, tenemos la zona metaestable, donde una solución aunque supersaturada nunca produciría núcleos cristalinos, salvo que se produzca una perturbación de las condiciones. En esta zona, la aplicación de

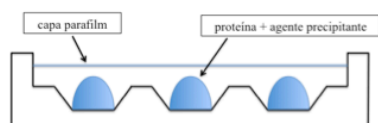
microcristales que actúen como base para la cristalización, en un método conocido como *microseeding*, puede generar la formación de cristales.

Existen tres técnicas principales para la obtención de cristales de proteína. En la figura 3.8 se muestra el esquema de cada una de ellas.

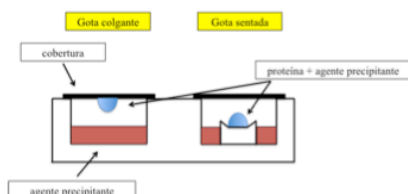
1. **Cristalización en lote.** En este caso el agente precipitante es añadido directamente a la solución de proteína, llevando la solución de forma instantánea a un estado de alta supersaturación (figura 3.8a).
2. **Difusión de vapor.** Este es el método más usado y el empleado por nosotros en todos los experimentos de cristalización. En este método, la solución de proteína se mezcla con una solución de precipitante y la mezcla se deja equilibrar por difusión de vapor frente a un reservorio que contiene un volumen mayor de solución de precipitante. Este procedimiento se puede realizar de dos formas diferentes. La primera de ellas es el método conocido como gota colgante (*the hanging drop method*), donde sobre una superficie de cristal siliconizada o de plástico antideslizante se mezcla un volumen pequeño (0.5-2 μL) de la solución de proteína a la concentración adecuada con, típicamente, el mismo volumen de la solución de precipitante. Una vez mezclada, la gota se coloca boca abajo sobre un reservorio que contiene la solución del precipitante (0.5-1 mL), sellando el pocillo con grasa para evitar la evaporización. La gota con proteína y la solución del pocillo tienden a equilibrarse lentamente, apareciendo con el tiempo cristales de proteína en la gota. Otra forma de aplicar el método y la empleada la mayoría de las veces por nosotros es la técnica de gota sentada (*the sitting drop method*) donde la gota que contiene la solución de proteína y de cristalización reposan sobre un estante o puente comunicado con un reservorio donde se añade la solución de cristalización. De nuevo la cámara, donde se encuentra el pocillo, es sellada para que se dé el equilibrio entre el pocillo y la gota y evitar la evaporización de las soluciones (figura 3.8b).
3. **Diálisis.** Este método consiste en realizar una diálisis con la muestra de proteína, usándose como tampón de diálisis el agente de precipitación (figura 3.8c). De nuevo en este caso existen dos variantes a la hora de llevar a cabo el experimento. En una de ellas se emplean botones de diálisis para evitar usar grandes volúmenes. En el segundo de los casos se emplea un capilar que se introduce en la solución del agente precipitante, donde se alcanza el equilibrio de las soluciones.

Por último, en cuanto a la cristalización de proteínas cabe mencionar de nuevo que es un proceso que depende de múltiples variables. Además de los parámetros mencionados inicialmente, (temperatura, concentración de proteína, la naturaleza y concentración del precipitante, el pH de la solución) existen distintos aditivos como metales, haluros, aminoácidos, azúcares, y muchos otros compuestos químicos, que se pueden añadir para aumentar las probabilidades de aparición de cristales. Dada la gran cantidad de variables implicadas, los experimentos de cristalización de una nueva proteína comienzan con cribados exhaustivos de multitud de condiciones de cristalización en formato de placa de 48 o 96 condiciones (Gavira, 2015).

a) Lote de cristalización



b) Difusión por vapor



c) Diálisis

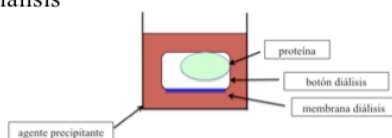


Figura 3.8. Técnicas empleadas en la obtención de cristales de proteína. Se representan esquemáticamente los tres métodos principales desarrollados para la formación de cristales de proteína, **a)** lote de cristalización, **b)** difusión de vapor y **c)** diálisis.

Por otro lado, en cuanto a uno de los parámetros modificables, existe una prueba que se emplea para evaluar que la concentración de proteína a cristalizar es correcta, o al menos útil para la realización de los cribados. Este experimento es el PCT (*Pre-Crystallization Test*), el cual nos va a permitir saber la concentración de proteína aproximada que vamos a emplear en los experimentos. Este experimento es realizado mediante la técnica de difusión de vapor, en concreto la de gota colgante. La concentración inicial recomendada suele ser de 5 mg/mL, aunque normalmente hay que aumentarla (www.hamptonresearch.com). Además el PCT también nos puede dar una orientación de que agente precipitante le viene mejor a la formación de cristales en cada caso. En la tabla 3.3 se muestra la composición de las 4 soluciones empleadas en esta prueba. Una vez sabemos que concentración queremos usar de proteína nos disponemos a realizar los procesos de cribado de condiciones mediante el método descrito como gota sentada (*the sitting drop method*) (figura 3.8).

Solución	Composición
A1	0,1M Tris-HCl pH 8,5, 2,0 M sulfato amónico
A2	0,1 M Tris-HCl pH 8,5, 0,2 M MgCl ₂ , 30% (w/v) PEG 4.000
B1	0,1 M Tris-HCl pH 8,5, 1,0 M sulfato amónico
B2	0,1 M Tris-HCl pH 8,5, 0,2 M MgCl ₂ , 15% (w/v) PEG 4.000

Tabla 3.3. Composición de las soluciones del PCT. Se muestran las cuatro soluciones empleadas para el analizar las potenciales condiciones de cristalización.

Muchas veces es de interés cristalizar una proteína con un ligando o un inhibidor potencial para entender mejor qué residuos de la proteína están involucrados en su función. Si el ligando de interés no se obtiene unido a la proteína directamente durante la expresión y purificación de la misma, se puede introducir en el experimento de cristalización mediante dos técnicas: el *soaking* (incubación) y la cocrystalización.

En el caso de la cocrystalización, se mezcla la proteína con el ligando deseado, para después mezclar el complejo con el agente precipitante. Del mismo modo que en lo explicado anteriormente, la cocrystalización puede verse afectada además por la concentración relativa de proteína y de ligando.

En cuanto a la técnica *soaking*, ésta consiste en añadir el ligando directamente a la gota donde se encuentran los cristales. Sin embargo hay que considerar varios aspectos debido a que los cristales son estructuras muy delicadas. Los cristales de proteína son frágiles y la adición de una solución diferente (en la que se encuentra el ligando) puede alterar la estructura de éstos rompiéndolos e incluso disolviéndolos. Además los tiempos de *soaking* y la concentración del ligando deben ser optimizados para conseguir un resultado óptimo (Hassell *et al.*, 2007).

3.5.2. Recogida y procesamiento de datos

Una vez detallado como obtener cristales de proteína, éstos pueden ser difractados en instalaciones especiales llamados sincrotrones. Los cristales de proteína son susceptibles de sufrir daños de radiación al ser expuestos a los rayos-X. Para minimizar esos daños los cristales son congelados en nitrógeno líquido. Dado la fragilidad de los mismos, los cristales deben ser incubados en unas condiciones especiales que los preserve y que implica el uso de crioprotectores. Los crioprotectores usados más habitualmente son el glicerol, azúcares como glucosa o sacarosa, MPD y PEG. En nuestro caso, normalmente con glicerol al 10-20% (v/v) se han obtenido buenos resultados. La exposición del cristal al crioprotectante y la extracción del cristal de la gota se realiza bajo el microscopio, para comprobar en todo momento el estado del cristal. Para la transferencia del cristal a los viales sumergidos en nitrógeno líquido, tras la incubación con el crioprotectante, los cristales deben ser pescados, empleando para ello un asa o lazo de nilón (material que no difracta frente a los rayos-X). Una vez pescado, el lazo con el cristal es introducido rápidamente en viales, los cuales ya están sumergidos en nitrógeno líquido y perfectamente identificados para poder ser difractados ya sea en una fuente de rayos-X local (*e.g.* un difractor de ánodo rotatorio) o almacenados para ser difractados en un sincrotrón.

El empleo de crioprotectores durante el proceso de congelación de los cristales tienen dos principales ventajas. La primera de ellas es que se reduce el daño de las radiaciones a las que se ve sometido el cristal durante la recolección de datos en los sincrotrones, las cuales son de alta intensidad. La segunda de ellas es la capacidad de conservar los cristales durante un tiempo, ya sea para el transporte o para su almacenamiento. Otra de las funciones de estas sustancias es que evitan la formación de hielo en los cristales durante su congelación con nitrógeno líquido ya que la solución vitrifica, evitando la aparición de grietas en éstos que disminuirían su calidad (Pflugrath, 2015; www.ccp4.ac.uk).

Los sincrotrones son grandes infraestructuras donde se genera potentes haces de rayos-X de gran intensidad y brillo y de pequeño tamaño (alta colimación). Los sincrotrones constan de un acelerador lineal, donde se genera un haz de electrones de alta energía, que es posteriormente acelerado hasta casi la velocidad de la luz en un acelerador circular (*booster ring*); en ese momento los electrones son inyectados en un anillo más grande (*storage ring* o anillo de almacenamiento), típicamente de 300 m de circunferencia, donde los electrones se mantienen en trayectorias aproximadamente circulares mediante la aplicación de fuertes campos magnéticos. Cada vez que el camino descrito por el haz de electrones se curva, se genera un intenso haz de rayos-X

que se emite hacia el exterior y en dirección tangencial a la trayectoria de los electrones. Este haz de rayos-X, denominado radiación de sincrotrón, es recogida y modificada en las estaciones de trabajo (*beamlines*) situadas a lo largo de la circunferencia del anillo de almacenamiento. Las estaciones de trabajo consisten de elementos ópticos para la recogida y focalización del haz de rayos-X sobre la muestra así como de cabinas de control para la colección y procesamiento de los datos generados durante la difracción de los cristales. Además del alto brillo y colimación del haz de rayos-X, otra ventaja de la radiación de sincrotrón en cristalografía es que la longitud de onda apta en el rango espectral puede ser seleccionada gracias a monocromadores. Esta propiedad es usada para la difracción anómala de longitud de onda múltiple (*Multiple-wavelength anomalous dispersion*), la cual aprovecha las diferencias en la intensidad de la difracción producidas por el fenómeno conocido como dispersión anómala, la cual tiene su base en la diferente composición de átomos que posee una proteína. Esta técnica es de gran utilidad en la resolución de estructuras proteicas. En la figura 3.9 se muestra una imagen del sincrotrón ALBA (Barcelona, España).

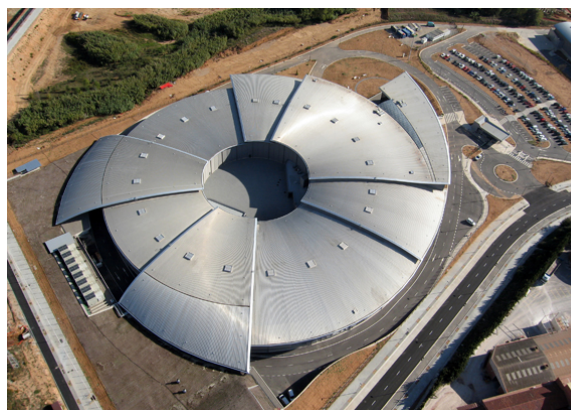


Figura 3.9. Sincrotrón ALBA de Barcelona. En la imagen se muestra una visión aérea del único sincrotrón de España, situado en la provincia de Barcelona.

El experimento de difracción (figura 3.10) requiere colocar el lazo que contiene el cristal vitrificado en un dispositivo especial llamado minidifractómetro. El minidifractómetro consiste en un goniómetro dispuesto en un brazo móvil, lo que permite posicionar y orientar el cristal en el espacio de la forma más adecuada. Una vez centrado, el cristal se expone durante breves intervalos de tiempo (*e.g.* 0,1-0,5 s) al haz de rayos-X al mismo tiempo que el cristal rota (*e.g.* 0,1-0,5°) a través de un eje perpendicular al haz. La rotación total del cristal depende del grupo espacial de simetría al que pertenezca. A mayor simetría, menor debe ser la rotación para obtener un conjunto de datos completo.

Los rayos-X difractados por el cristal son recogidos sobre un detector, de manera que cada exposición se traduce en una imagen de difracción de rayos-X con patrones

característicos de manchas de difracción (*diffraction spots*) que se originan por la interferencia constructiva de los rayos-X difractados (figura 3.10). Cada mancha de difracción se corresponde con una reflexión específica y se caracteriza por su posición en el detector y su intensidad. Cada reflexión tiene asociado un índice de Miller (hkl), que se corresponde con la familia de planos cristalinos que difracta los rayos-X que la generan.

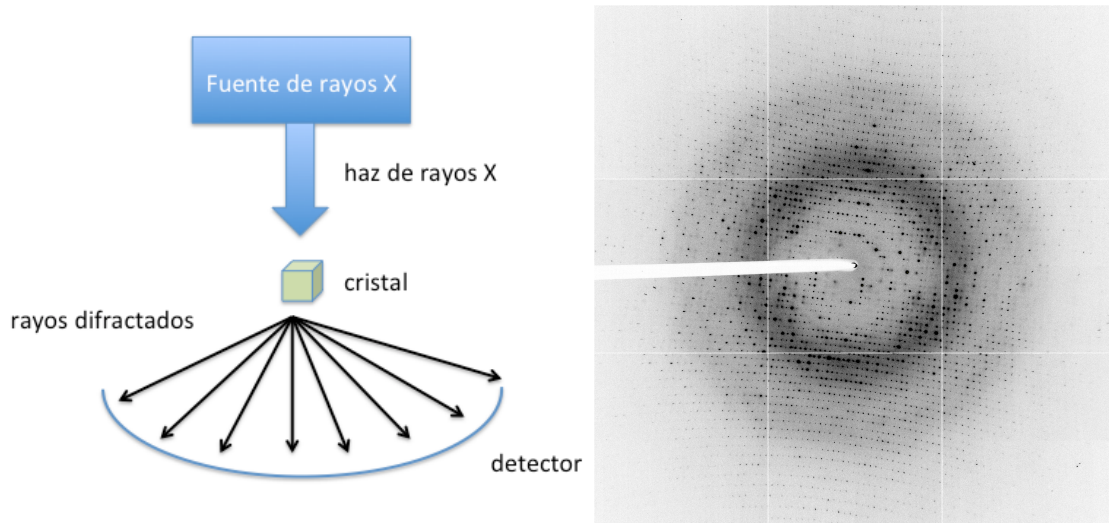


Figura 3.10. Componentes básicos de una determinación cristalográfica por rayos-X y patrón de difracción. A la izquierda se representa de forma básica la difracción de rayos-X de cristales de proteína. Una vez difractado el cristal de proteína mediante el empleo de rayos-X, los rayos difractados originan una imagen de difracción (parte derecha de la figura) en el detector. A partir de estos patrones es posible dilucidar la estructura tridimensional de la proteína difractada.

El procesamiento de datos de cristalografía consiste en la extracción de la intensidad e índice de Miller de cada reflexión a partir de una colección de imágenes de difracción. El procesamiento consiste en varios pasos. El primero es el indexado, que consiste en asignar una reflexión determinada a cada mancha de difracción y calcular la celda unidad y el grupo de simetría del cristal. Hay 65 grupos espaciales quirales en los que las proteínas pueden cristalizar, clasificados en 7 sistemas cristalinos en función de los elementos de simetría que contienen (tabla 3.4).

Sistema cristalino	Elementos	Restricciones
Triclínico	1 centro de simetría o ninguna	$a \neq b \neq c; \alpha \neq \beta \neq \gamma$
Monoclínico	1 eje binario o 1 plano de simetría	$a \neq b \neq c; \alpha = \gamma = 90^\circ \beta \neq 90^\circ$
Ortorrómbico	3 ejes binarios o 3 planos de simetría	$a \neq b \neq c; \alpha = \beta = \gamma = 90^\circ$
Tetragonal	1 eje cuaternario	$a=b \neq c; \alpha = \beta = \gamma = 90^\circ$
Trigonal	1 eje ternario	$a=b=c; \alpha = \beta = \gamma \neq 90^\circ$
Hexagonal	1 eje senario	$a=b \neq c; \alpha = \beta = 90^\circ \gamma = 120^\circ$
Cúbico	4 ejes ternarios	$a=b=c; \alpha = \beta = \gamma = 90^\circ$

Tabla 3.4. Clasificación de los sistemas cristalinos en función de los elementos de simetría.

Se muestran los sistemas cristalinos, así como sus simetrías y las restricción en cuanto a los ángulos del cristal.

A continuación, los parámetros de la celda unidad son refinados u optimizados. El segundo paso se denomina integración. Durante la integración, se mide la intensidad de cada una de las reflexiones listadas en el indexado al tiempo que se optimizan diversos parámetros de la geometría del experimento de difracción (*e.g.* la distancia cristal-detector) y propiedades del cristal (*e.g.* mosaicidad).

Uno de los programas usados con este fin es MOSFLM (Battye *et al.*, 2011). El uso de este tipo de programas también es de alto valor a la hora de identificar ciertos problemas, como pueden ser manchas muy débiles, fraccionadas o de baja calidad en la imagen de difracción, mosaicidad muy alta del cristal, difracción anisotrópica, presencia de manchas de difracción o la presencia de anillos debido al hielo, entre otros (Battye *et al.*, 2011).

Finalmente, en el tercer paso, se realiza un escalado y promediado de la intensidad de todas las reflexiones relacionadas por simetría y se genera una lista de reflexiones únicas, intensidades y desviaciones estándar asociadas. Cada valor de intensidad, $I(h,k,l)$, es convertido en el modulo de un Factor de estructura, $F(h,k,l)$, que es proporcional a la raíz cuadrada de la intensidad correspondiente. La calidad del escalado de los datos de difracción se valora mediante el cálculo de diversos parámetros y estadísticos, que se analizan tanto globalmente como por capas de resolución. Los más importantes son: la completitud, la multiplicidad/redundancia, el promedio de la intensidad sobre la desviación estándar $[I / \sigma(I)]$ y varios estadísticos que intentan capturar cómo de similares son las intensidades de las reflexiones relacionadas por simetría (*e.g.* R_{merge} , R_{meas}).

La completitud es el porcentaje de todas las reflexiones cuyas intensidades fueron medidas experimentalmente con respecto al número total de reflexiones posibles, calculadas a partir de la celda unidad y el grupo espacial de simetría. Idealmente, la completitud global y en la capa de mayor resolución debería ser 100%, aunque en la práctica se aceptan valores de 90-95%.

El número de medidas independientes de cada reflexión es su multiplicidad o redundancia. La multiplicidad depende de la simetría cristalina y del rango angular de la recogida de datos. En general, se aspira a que la multiplicidad del conjunto de datos sea lo más alta posible ya que aumenta la significación estadística de las medidas de intensidad.

La $I/\sigma(I)$ es equivalente a la razón de la señal-ruido y, por tanto, describe la significación estadística de las medidas de intensidad. Esta proporción disminuye a medida que aumenta la resolución. Tradicionalmente, la resolución nominal del conjunto de datos se asignaba a aquella capa de resolución en la que $I/\sigma(I)$ era 2,0, aunque en la actualidad se aceptan reflexiones con valores menores de $I/\sigma(I) > 1,0$.

El R_{merge} y R_{meas} miden la dispersión de n medidas independientes de la intensidad de las reflexiones ($I_i(hkl)$) alrededor de sus promedio ($\bar{I}(hkl)$) (figura 3.11). El R_{meas} se diferencia del R_{merge} en que el R_{meas} tiene en cuenta la multiplicidad de cada reflexión.

$$R_{merge} = \frac{\sum_{hkl}^n \sum_{i=1}^n |I_i(hkl) - \bar{I}(hkl)|}{\sum_{hkl}^n \sum_{i=1}^n I_i(hkl)}$$

$$R_{meas} = \frac{\sum_{hkl} (n/n-1)^{1/2} \sum_i |I_i(hkl) - \langle I_i(hkl) \rangle|}{\sum \sum_i I_i(hkl)}$$

Figura 3.11. Expresiones de los indicadores estadísticos empleados en cristalografía de rayos-X de proteínas.

En adición a R_{merge} y R_{meas} se han empezado a emplear otros estadísticos como el coeficiente de correlación de Pearson ($CC_{1/2}$), calculado a partir de dos mitades de

grupos de datos de igual tamaño y elegidos al azar. Este parámetro da una valoración directa de las proporciones relativas para cada señal y ruido que contribuye a la variación de los datos en una resolución dada. La resolución nominal del conjunto de datos se puede fijar de acuerdo con el $CC_{1/2}$ en aquella capa de resolución donde $CC_{1/2} = 0,25$.

Por último, otro parámetro a tener en cuenta es el coeficiente CC^* , que es el coeficiente de correlación entre un grupo de datos completos de las intensidades reales y las estimadas gracias a la expresión de la figura 3.12. El programa de escalado usualmente empleado durante el presente trabajo es SCALA (Evans, 2011) y su versión actualizada, Aimless (Evans y Murshudov, XXX), de la suite de programas CCP4.

$$CC^* = \sqrt{\frac{2CC_{1/2}}{1 + CC_{1/2}}}$$

Figura 3.12. Expresión derivada del coeficiente de correlación de Pearson. A través de esta expresión es posible valorar la calidad de los datos, además de permitir la comparación directa de la calidad del modelo cristalográfico y la calidad de los datos en una misma escala.

3.5.3. Problema de las fases

Una vez difractados los cristales se obtienen las imágenes de difracción (figura 3.10, derecha) a partir de las cuales es necesario resolver el problema de las fases para poder llegar a determinar la estructura tridimensional de la proteína. La transformada de Fourier (figura 3.13) es la solución matemática que nos permite relacionar la densidad electrónica de una proteína, es decir, donde están localizados cada uno de sus átomos, con la información obtenida de las imágenes de difracción, la cual contiene los índices de las reflexiones y las intensidades de los haces difractados. Además, es también necesario conocer la diferencia de fase entre las ondas originadas tras la difracción, dato no reflejado en las imágenes de difracción de cristales macromoleculares. Este último concepto es conocido como el problema de fases.

$$\rho(xyz) = 1/V \sum_{hkl} |F(hkl)| e^{-2\pi i[hx + ky + lz - \phi(hkl)]}$$

Figura 3.13. Expresión de la transformada de Fourier que representa la función de densidad electrónica. Expresión necesaria para la resolver el problema de fases y por lo tanto para la determinación de la estructura tridimensional de las proteínas estudiadas por cristalografía de rayos-X. $\rho(xyz)$ representa la densidad electrónica de cada átomo en una coordenada determinada; V se refiere al volumen de la celda unidad, repetición periódica mínima que forma el cristal; $F(hkl)$ son los factores estructurales, representantes de las ondas resultantes de la dispersión de todos los átomos (sus módulos se encuentran directamente relacionados con las intensidades de las reflexiones); $\phi(hkl)$ representa las fases que se quieren calcular.

De esta forma la transformada de Fourier (figura 3.13) relaciona el espacio real (posicionamiento de los átomos) representado por la densidad electrónica ($\rho(xyz)$), con el espacio recíproco, que se manifiesta en el patrón de difracción. A la hora de resolver una estructura macromolecular mediante cristalografía de rayos-X existen tres métodos principales: (i) método de Reemplazamiento Isomorfo Múltiple (MIR), (ii) método de Difracción Anómala Múltiple (MAD), donde también encontramos la Difracción Anómala a Longitud Única (SAD) y (iii) Reemplazamiento molecular (MR). En nuestro caso se empleó éste último y SAD. Existen otros métodos cuya aplicación es más restringida.

3.5.3.1. Reemplazamiento Molecular (MR)

En el método MR, las fases (conocidas) del cristal de una proteína homóloga de estructura existente son empleadas como primera aproximación para el cálculo de las fases (desconocidas) para un nuevo cristal de proteína. La estructura homóloga se denomina modelo de MR. Normalmente, se necesita que la identidad de secuencia entre el modelo y la nueva proteína sea >45%. En ocasiones, el MR puede funcionar correctamente con modelos cuya identidad de secuencia es menor del 45% y/o con modelos parciales (e.g. cuando el nuevo cristal contiene dos proteínas diferentes y solo

existe modelo para una de ellas). En general, el éxito del MR con un modelo determinado depende de los siguientes factores: (a) Homología; cuanto mayor sea la identidad de secuencia más fácil resultará la resolución de la estructura; (b) Integridad; el modelo debe representar una fracción significativa de la estructura desconocida; y (c) Flexibilidad; algunas proteínas, como en el caso de anticuerpos, poseen dominios que adoptan diferentes conformaciones debido a su alto grado de flexibilidad, lo cual dificulta la aplicación de MR en estas zonas debido a que la orientación relativa de los dominios es variable (Scapin, 2013).

A continuación, se emplean procedimientos computacionales para colocar el modelo de MR en la celda cristalina del nuevo cristal y calcular la orientación y posición más probable para cada una de las moléculas presentes en la unidad asimétrica. La orientación se calcula en primer lugar mediante el uso de la función de rotación, y a continuación se localiza la posición mediante la función de traslación. La función de traslación se aplica iterativamente a cada una de las moléculas que es preciso localizar, cada una de las cuales podría tener una función de rotación diferente.

Para la determinación de la función de rotación en el MR (o de la subestructura de átomos pesados introducidos en el cristal de proteína en los métodos de difracción anómala), se emplea la función o ecuación de Patterson. La ecuación de Patterson (ecuación 3.1) suprime la información de las fases ($\phi(hkl)$) de la expresión de la figura 3.13, a la vez que proporciona información de los vectores de la posición de los distintos átomos de la estructura ($P(uvw)$). Calculada esta función se obtiene la posición absoluta de dichos átomos, lo que nos va a permitir en último término la visualización de la densidad electrónica y por lo tanto de la proteína a nivel atómico.

$$P(uvw) = \frac{1}{V} \sum_{hkl}^{+\infty} |F(hkl)|^2 \cos 2\pi(hu + kv + lw)$$

Ecuación 3.1. Ecuación de Patterson. Ecuación simplificada y derivada de la expresión de la transformada de Fourier, la cual suprime la información de las fases y añade $P(uvw)$, representando la posición absoluta de cada átomo, que finalmente permitirá calcular las fases y por lo tanto la densidad electrónica en cada átomo de la proteína.

Una vez encontrada la rotación, y si existe simetría translacional, se procede a trasladar el modelo en la orientación adecuada hasta encontrar la posición en el cristal.

Este proceso se repite hasta posicionar todas las copias de las macromoléculas presentes en la unidad asimétrica. En el caso de que haya múltiples copias de una proteína en la unidad asimétrica, se habla de la existencia de simetría no cristalográfica (NCS, *non-crystallographic symmetry*).

En el caso de que el nuevo cristal sea de la misma proteína que el modelo, y tenga el mismo grupo espacial y, aproximadamente, la misma celda unidad (más correctamente, que el nuevo cristal sea isomórfico con el cristal modelo), se puede resolver su estructura mediante refinamiento de cuerpo rígido (*Rigid Body Refinement*) sin necesidad de realizar búsquedas rotaciones y traslacionales.

3.5.3.2. SAD (*Single-Wavelength Anomalous Dispersion*)

Otra de las técnicas empleadas para la resolución estructural de proteínas es SAD. Esta técnica se basa en el fenómeno de difracción anómala, que sucede cuando determinados átomos poseen longitudes de onda de absorción muy próximas a la longitud de onda de los rayos-X del experimento de difracción. Dichos átomos pueden ser constituyentes habituales de las macromoléculas biológicas (*e.g.* S, P, Fe) o átomos que se incorporan en los cristales solo como herramientas para determinar estructuras para las que no existen modelos en el PDB. Se y Hg son dos átomos de este último tipo. En situaciones normales, la quiralidad de las moléculas de proteína asegura que (h,k,l) sea siempre equivalente por simetría a $(-h,-k,-l)$; la equivalencia de las intensidades de ambas reflexiones se denomina ley de Friedel. Sin embargo, la presencia en el cristal de átomos capaces de experimentar difracción anómala provoca la ruptura de la ley de Friedel, porque la combinación de la absorción con la difracción causa que la intensidad de cada reflexión (h,k,l) sea ahora diferente de la de su par de Friedel $(-h,-k,-l)$. En la mayor parte de los casos esta diferencia de intensidad es muy pequeña y es difícil medir la diferencia entre $I(h,k,l)$ e $I(-h,-k,-l)$, que da una medida de la señal anómala. Para poder medirla con precisión, se han de incorporar muchos átomos con difracción anómala en las macromoléculas y/o seleccionar la longitud de onda de los rayos-X para maximizar el efecto de difracción anómala.

Aunque convencionalmente la señal anómala se mide a partir de varios conjuntos de datos, cada uno de ellos a una longitud de onda diferente, en muchos casos un solo conjunto de datos recogido a la longitud de onda de máxima absorción (el “pico” de absorción) puede contener suficiente información para resolver una estructura cristalográfica completa (Dodson, 2003).

3.5.3.3. Refinamiento y validación

Después de orientar y posicionar el modelo de MR en el cristal, comienza el proceso de construcción del modelo atómico o trazado de la nueva estructura. Para ello, se calculan mapas de densidad electrónica a partir de los factores de estructura observados experimentalmente y los parámetros del modelo (posiciones atómicas (x,y,z) y factores B). Estos últimos se utilizan para calcular un primer conjunto de fases aproximadas para la estructura. Los mapas de densidad revelan aquellas regiones donde el modelo inicial no representa correctamente la nueva estructura y que deben ser modificadas.

La corrección sistemática de todos los errores implica la iteración entre una fase de construcción automática o manual mediante el uso de programas gráficos, en la que los átomos del modelo de MR son ajustados a las posiciones indicadas en el mapa de densidad electrónica, y una fase de refinamiento.

El refinamiento es un procedimiento matemático en el que se optimizan los parámetros del modelo (el ajuste del modelo trazado) de acuerdo con los datos experimentales. Para la optimización se aplica una función matemática al conjunto de parámetros a refinar que puede ser de mínimos cuadrados (*least-square refinement*) o de máxima verosimilitud (*maximum likelihood*). Salvo en aplicaciones específicas, el refinamiento de estructuras macromoleculares se realiza mediante máxima verosimilitud. Cuando solo se varían la posición y orientación del centro de masas de las proteínas (por ejemplo, un cuerpo rígido por cadena o varios cuerpos rígidos, en el caso de que haya diversos dominios por cadena polipeptídica), pero no la posición de los átomos individualmente, se considera un refinamiento de cuerpo rígido, especialmente útil durante la etapa de MR o en etapas tempranas de la resolución estructural. En el caso de que haya múltiples copias de una proteína en la unidad asimétrica, esta redundancia se puede utilizar en forma de restricciones de simetría no cristalográfica, lo que conlleva una reducción efectiva de los parámetros ajustables.

A lo largo del refinamiento se comparan los factores estructurales calculados a partir del modelo con los factores estructurales experimentales, tendiendo en cada uno de los ciclos a minimizar esas diferencias hasta un mínimo. Dichas diferencias se conceptualizan en los *R factors* (factores R), el *R factor* de trabajo o cristalográfico (*R-work*) y el *R-free*. En la figura 3.14 se muestran las ecuaciones de estos dos factores. Para el cálculo del *R-free*, se separa una fracción de los datos experimentales (el *test set*, típicamente 5-10% de todas las reflexiones) que nunca se utilizan para el refinamiento. Con este procedimiento estadístico de validación cruzada se evita introducir sesgos durante el refinamiento y construcción del modelo. En cada ciclo de refinamiento – en

el que se empleó el programa phenix.refine (Afonine *et al.*, 2012) de la suite de programas de cristalografía PHENIX – se hizo un seguimiento del mapa de densidad electrónica con la construcción manual y/o automática del modelo cristalográfico. El modelo cristalográfico se completa con la adición de ligandos/cofactores, moléculas de agua, iones y cualquier otro componente del solvente, de la condición de cristalización y/o del crioprotectante. La posición de estas moléculas se puede hacer automáticamente por los programas de refinamiento, con ajustes manuales posteriores.

$$R = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}}$$

Figura 3.14. Expresión matemática de los *R*-factors (*R*-work y *R*-free). *R* representa el *R*-factor, mientras que F_{hkl} es el factor de estructura para las reflexiones *hkl*, parámetro que describe como la materia dispersa la radiación incidente (rayos-X en nuestro caso). Se estima en base a las reflexiones observadas (obs) y calculadas (calc).

Al final del refinamiento es necesario validar la estructura con el fin de garantizar la calidad del modelo obtenido. Para ello, se analiza exhaustivamente las características geométricas y estereoquímicas de la estructura y se corrige cualquier desviación significativa de los valores esperados. Por ejemplo, se analiza que las distancias interatómicas obtenidas y los ángulos de enlace estén dentro de unos criterios aceptables. Las medidas de la desviación de las distancias (en Å) y ángulos (en °) de enlace se calculan mediante la desviación cuadrática media (RMSD, *Root-Mean-Square Deviation*) frente a valores de referencia de moléculas pequeñas. También se verifica que los ángulos de torsión del enlace peptídico (Φ y ψ) sean energéticamente favorables y que no existan choques estéricos. Este tipo de validaciones son fácilmente comprobadas gracias al mapa de Ramachandran, el cual muestra conformaciones estables, así como las prohibidas, de los ángulos de torsión de los distintos aminoácidos, las cuales no podrían ser posibles. En la figura 3.15 se muestra el mapa de Ramachandran, mostrando las combinaciones permitidas y prohibidas de los ángulos Φ y ψ del enlace peptídico. Del mismo modo, los valores de los factores *B* asociados a cada átomo o grupos de átomos de la estructura deben de presentar valores físicamente aceptables.

Estos valores por lo general son menores en las zonas internas de la proteína y mayores en las zonas externas, donde existe un contacto directo con el solvente (Drenth, 2007).

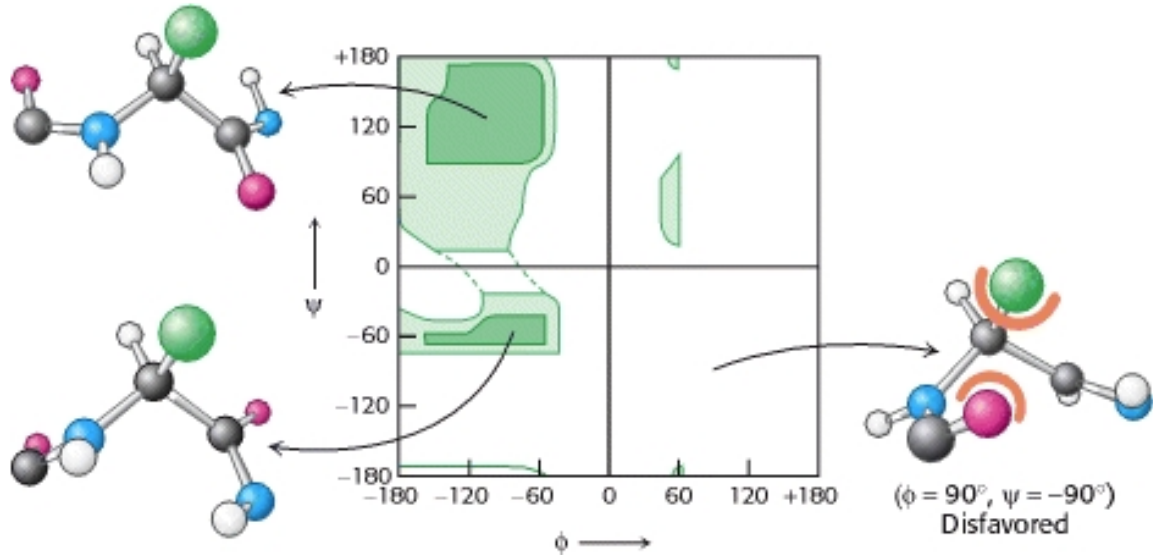


Figura 3.15. Mapa de Ramachandran mostrando las combinaciones permitidas y prohibidas de Φ (Phi) y ψ (Psi) en el enlace peptídico. No todas las combinaciones de los ángulos son permitidas (estructura derecha), debido a la colisión estérica. Las zonas más favorables se muestran en verde oscuro. Las zonas que se encuentran al límite de la permisividad se muestran en verde claro. Imagen extraída de Berg *et al.*, 2002.

Resultados

4. Resultados

4.1. Cisteína desulfurasa del sistema CSD (CsdA)

4.1.1. Clonaje de CsdA

El gen codificante para la CsdA se obtuvo del genoma de *E. coli* (XL1blue) mediante la técnica PCR con oligonucleótidos específicos para este gen (anexo I, tabla I). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes, se clonó en el vector de expresión, con resistencia a kanamicina, pETM11. En este caso, el vector y el inserto no fueron digeridos con las mismas enzimas, siendo el vector digerido con las enzimas de restricción *NcoI/BamHI* y el inserto con *BsaI/BamHI*. Una vez obtenida la construcción pETM11-CsdA, se transformó en la cepa de clonaje de *E. coli* XL1blue, creciendo colonias en placas de LB con agar al 1,5% y kanamicina 50 µg/mL. Al comprobar que la construcción era correcta mediante PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen (anexo I, tabla I), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con kanamicina.

4.1.2. Expresión de CsdA

Para ello se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Esta prueba se realizó en 1 mL de LB más kanamicina a la concentración mencionada anteriormente. De esta forma las colonias fueron crecidas (310 K 800 rpm) hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación se disminuye la temperatura hasta 303 K y se induce el cultivo con la adición de 0,5 mM IPTG (3 horas 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso, el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonificado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 µL del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más kanamicina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin embargo, el tiempo de expresión se alargó a 12 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

4.1.3. Purificación de CsdA

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular generado fue resuspendido en el tampón de lisis (50 mM Tris-HCl pH 8,0, 500 mM NaCl, 10 mM imidazol y 2 mM β -mercaptoetanol), al que se le añade 10 mM de PMSF y 2 mM de PLP. Una vez resuspendido, éste es sonificado (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. A continuación el lisado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 μ m de diámetro de poro del material PES (Polietersulfona), el cual fue empleado en todos los casos de purificación en esta tesis. Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de un ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución de la proteína aumentando gradualmente la concentración de imidazol hasta 250 mM, respetando el resto de las concentraciones de los componentes del tampón. La proteína eluyó a 125 mM de imidazol (figura 4.1). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 45 kDa, correspondiente a la masa molecular del monómero de CsdA (figura 4.1).

A continuación las alícuotas de mayor pureza en el paso anterior son seleccionadas y dializadas (12 horas 277 K) en el tampón 50 mM Tris-HCl pH 8,0, 150 mM NaCl y 2 mM β -mercaptoetanol, añadiendo proteasa TEV (*Tobacco Etch Virus*) con el fin de eliminar la cola de histidinas. Seguidamente se repite el proceso de la IMAC con los mismos tampones en una purificación sustractiva, sin la aplicación de ningún gradiente, donde la proteína recombinante no se une a la columna por la eliminación de la cola de histidinas. De nuevo el resultado es analizado mediante PAGE-SDS (figura 4.2).

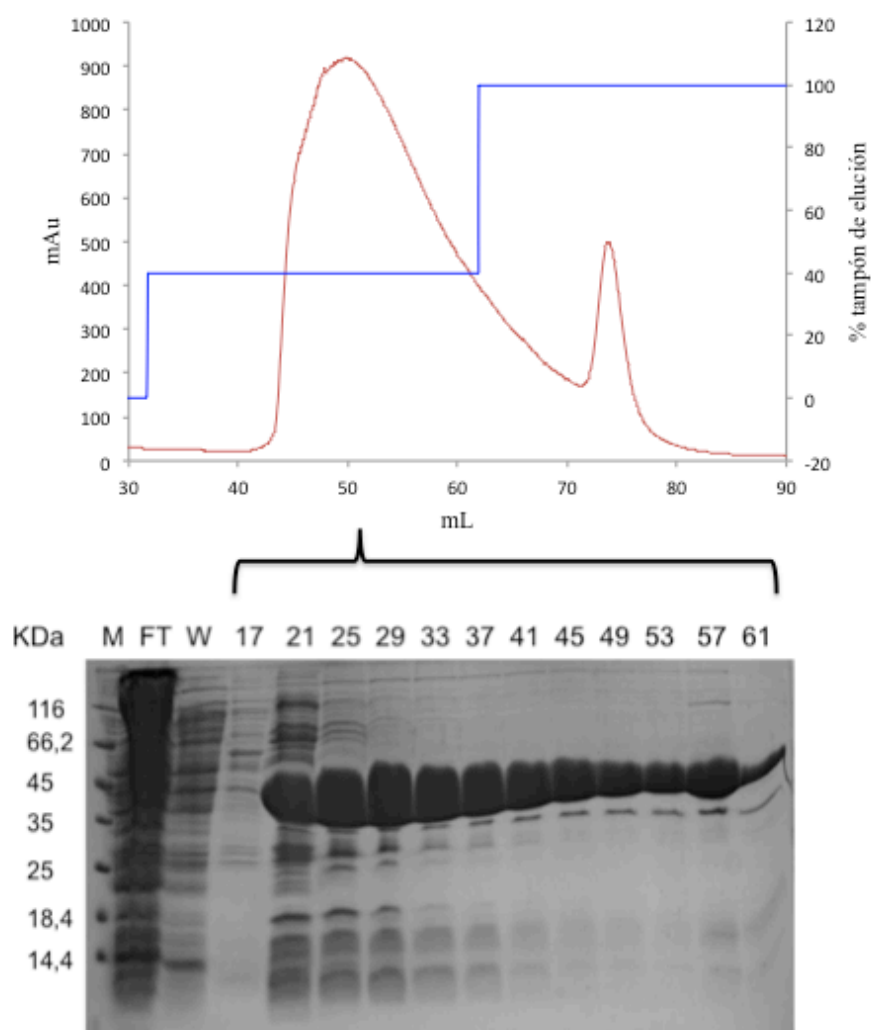


Figura 4.1. Purificación mediante IMAC de CsdA. Se muestra el cromatograma (arriba) y el gel de PAGE-SDS (abajo) resultantes. En el primero se representan las miliunidades de absorbancia (mAU) a 280 nm (línea roja) frente al volumen de elución en el cromatograma. La línea azul representa la concentración del tampón de elución. En la PAGE-SDS se muestran las alícuotas de la purificación, observándose que la mayoría de la proteína, sobre 45 kDa, corresponde a la CsdA recombinante. En el gel de PAGE-SDS M es el marcador, FT es el *flow-through*, W es el lavado y los números corresponden a las alícuotas analizadas.

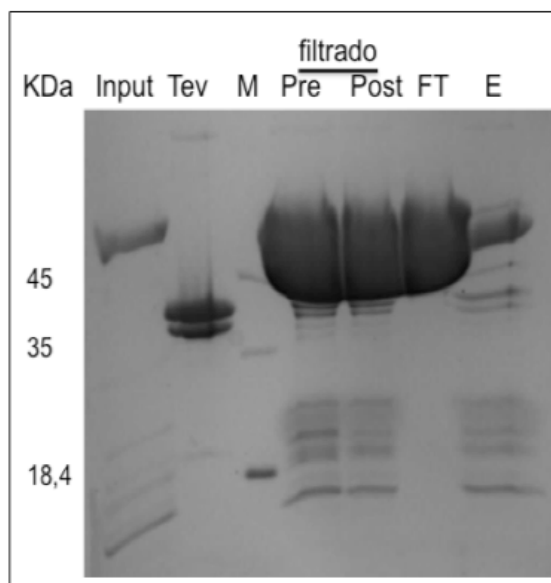


Figura 4.2. Purificación sustractiva de CsdA tras el corte con TEV. Se observa como la mayoría de la proteína recombinante está en el FT, es decir, ha sido cortada gracias a la acción de la proteasa. La muestra antes de ser cargada en la columna se filtró para descartar precipitantes, de ahí el análisis antes y después de filtrar, con filtros de tipo PES. El *Input* corresponde a la proteína antes de ser dializada y cortada. TEV corresponde al control donde solo se aplicó la proteasa, M es el marcador, pre y post (filtrado) son las muestras antes y después de filtrar la muestra tras la diálisis, respectivamente, FT es el *flow-through* y E se refiere a la elución.

Se puede observar en la figura 4.2 que la proteína recombinante (banda sobre los 45 kDa) ha perdido la cola de histidinas, pues es incapaz de unirse a la columna a bajas concentraciones de imidazol, al encontrarse en el FT (*flow-through*). Se pueden observar restos de proteína en la elución (E), por lo que se deduce que no toda la proteína se procesó de forma completa.

A continuación la proteína procedente del FT y lavado de la figura 4.2 es concentrada mediante el uso de concentradores *Millipore*, para ser cargada en una columna de filtración en gel (HiLoad 16/60 Superdex200, *GE Healthcare*), equilibrada previamente con el tampón 50 mM Tris-HCl pH 8,0, 150 mM NaCl y 2 mM β -mercaptoetanol. La filtración en gel fue llevada a cabo a 1 mL/min, eluyendo la proteína al volumen de elución esperado para el homodímero, 93,52 mL (figura 4.3), de aproximadamente 90 kDa. De nuevo las alícuotas del pico de elución obtenido fueron analizadas mediante PAGE-SDS, donde se puede apreciar el grado de pureza de la proteína.

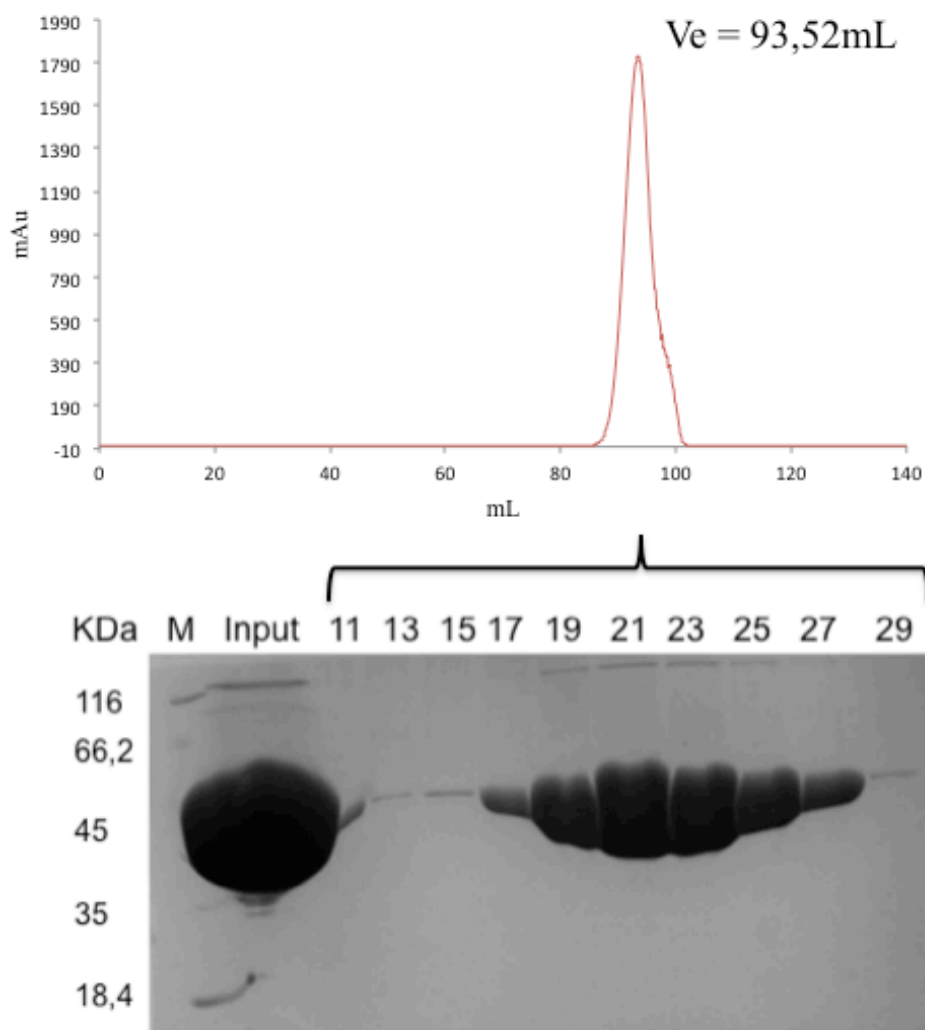


Figura 4.3. Purificación mediante filtración en gel de CsdA. Se representan el cromatograma (arriba) y el gel PAGE-SDS (abajo) correspondiente. Se puede observar que el volumen de elución (V_e) es de 93,52 mL correspondiente al dímero de la proteína. Se representa en rojo las miliunidades de absorbancia (mAU) a 280 nm. En la PAGE-SDS se observa la proteína con un gran grado de pureza. En el gel de PAGE-SDS M es el marcador, *Input* la muestra cargada en la columna y los números son las alícuotas analizadas.

4.1.4. Clonaje de los mutantes de CsdA (H120A y K250A)

Nota: el análisis estructural previo del sitio activo de CsdA sugirió que los residuos His120 y Lys250 de esta proteína podrían estar implicados en la transferencia de S a

Resultados

CsdE. Para comprender mejor este proceso, generamos dichos mutantes para su análisis estructural y funcional.

Los clonajes de los mutantes CsdA H120A y CsdA K250A se realizaron mediante los métodos de empalme-extensión y de *Quik Change*, respectivamente. En ambos casos se empleó la construcción de pETM11-CsdA como molde. En la tabla 1 del anexo I se muestran los oligonucleótidos empleados durante la mutagénesis. En la figura 4.4 se observa el gel de agarosa con la amplificación de los fragmentos generados durante el proceso de empalme-extensión de CsdA H120A.

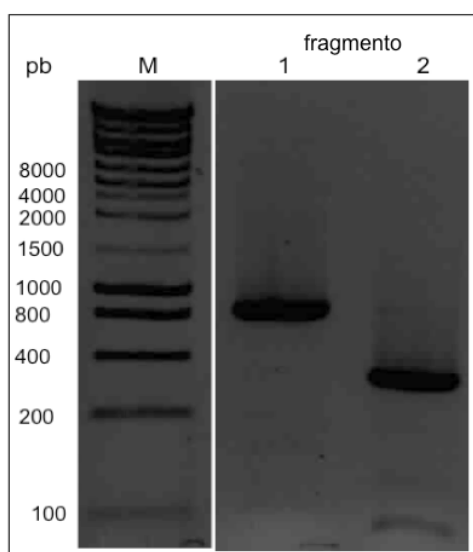


Figura 4.4. Análisis mediante gel de agarosa 1% de las PCR de los dos fragmentos de CsdA H120A mediante el método empalme-extensión. Se pueden observar los dos fragmentos del gen de CsdA portadores de la mutación diana (H120A), que fusionados formaran el gen completo de CsdA con la mutación deseada. pb son los pares de bases y M los marcadores.

Los fragmentos generados en la figura 4.4 fueron empleados como molde en una nueva reacción de PCR, con los oligonucleótidos específicos del gen codificante para CsdA (anexo I, tabla 1), para fusionarlos y seguir el procedimiento normal de clonaje, es decir, digerir el inserto de PCR resultante y ligarlo en el vector de expresión pETM11 para seguir los pasos indicados en materiales y métodos en cuanto a clonaje (apartado 3.2.1).

El segundo de los mutantes, CsdA K250A, fue generado en un solo paso de PCR mediante el método de *Quik Change*, empleando los oligonucleótidos correspondientes de la tabla 1 del anexo I. En este caso se emplearon diferentes cantidades del molde, de 5 a 50 ng. En la figura 4.5 se muestra el gel de agarosa empleado para analizar las muestras.

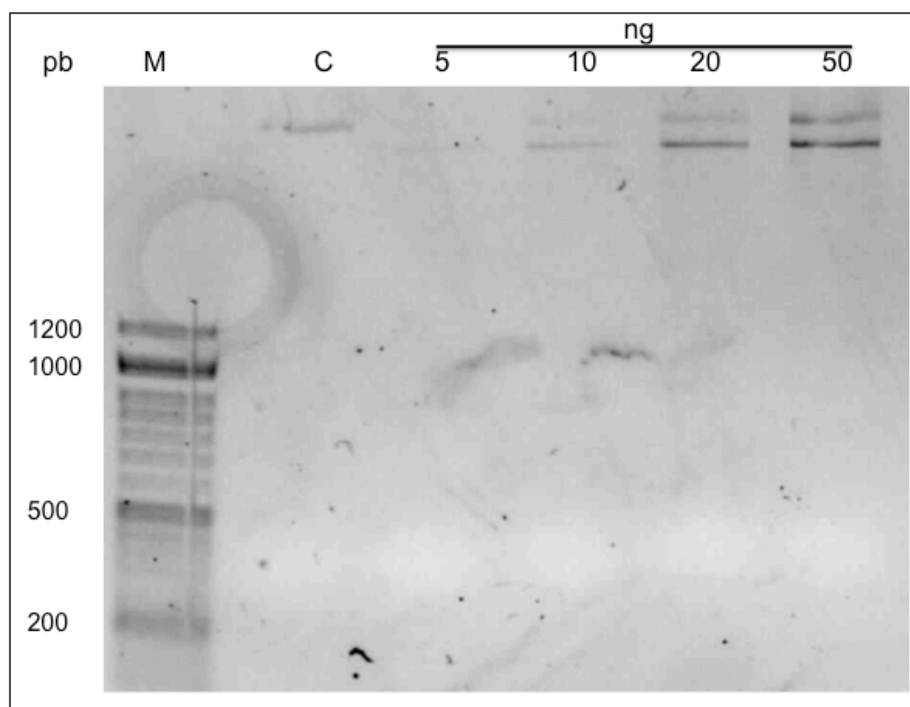


Figura 4.5. Análisis mediante gel de agarosa 1% de la PCR correspondiente a la mutagénesis dirigida mediante el método de *QuikChange* para la obtención del mutante K250A de CsdA. Como se puede observar en todos los casos (con diferentes cantidades de molde) parece haber funcionado la PCR apareciendo una banda a distinta altura que en el control negativo, siendo la amplificación de la construcción con la mutación. M son los marcadores, pb se refiere a los pares de bases y C es el control de carga.

Como se puede observar en la figura 4.5 en todos los casos existe amplificación. Las muestras fueron digeridas (enzima *DpnI*) y transformadas en la cepa *XL1blue* según lo indicado en el apartado correspondiente de materiales y métodos (apartado 3.2.1).

Tras la obtención de las construcciones con la mutación deseada introducida y su comprobación por secuenciación, se llevaron a cabo los mismos controles que en el caso de la proteína nativa hasta la prueba de expresión, donde se comportó de igual manera. El escalado se realizó en las mismas condiciones que la proteína nativa.

4.1.5. Expresión y purificación de los mutantes de CsdA (H120A y K250A)

La expresión de los mutantes de CsdA se realizó de un modo similar al empleado para la expresión de la proteína nativa (véase apartado 4.1.2).

Resultados

El mutante CsdA H120A se purificó llevando a cabo los mismos pasos que en el caso de la proteína nativa. Para ello se realizó primeramente una IMAC empleando las mismas condiciones que en la proteína nativa y los mismos tampones. En la figura 4.6 se muestra el resultado de la cromatografía de afinidad. Las muestras obtenidas en la elución fueron analizadas por PAGE-SDS, observándose la proteína a la altura del marcador de 45 kDa (figura 4.6).

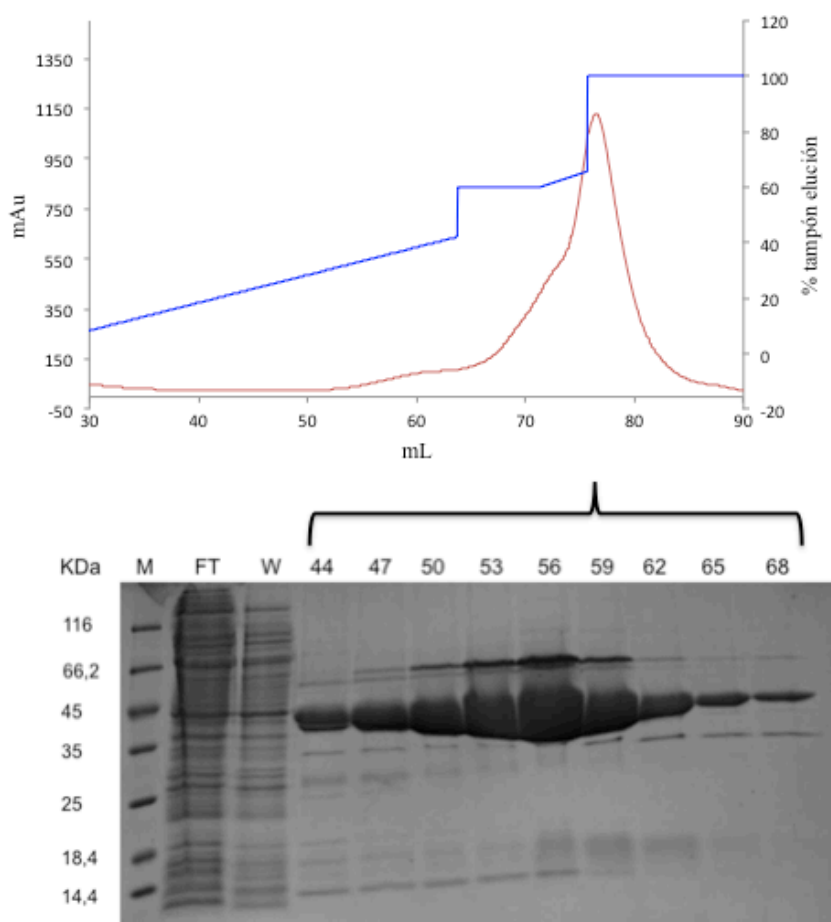


Figura 4.6. Purificación mediante IMAC de CsdA H120A. En el cromatograma (arriba) se representan las miliunidades de absorbancia (mAU) a 280 nm (línea roja) frente al volumen de elución en el cromatograma. La línea azul representa la concentración del tampón de elución. En la imagen del gel PAGE-SDS (abajo) se muestran las muestras de la purificación, observándose que la mayoría de la proteína, sobre 45 kDa, corresponde al mutante. En el gel de PAGE-SDS M indica los marcadores, FT el *flow-through*, W el lavado y los números corresponden a las alícuotas analizadas.

Una vez seleccionadas las alícuotas en la elución de la IMAC, éstas fueron concentradas mediante el uso de concentrados *Millipore* y cargada en la columna de filtración en gel Superdex200 10300 GL (*GE Healthcare*) empleando el mismo tampón que en el caso de la proteína nativa (50 mM Tris-HCl pH 8,0, 150 mM NaCl y 2 mM β -mercaptoetanol). En este caso no se realizó el procesamiento con la proteasa TEV. El flujo empleado en la filtración en gel fue de 0,5 mL/min. En la figura 4.7 se muestra el resultado de este paso, donde se observa un volumen de elución de 14,58 mL, correspondiendo en este tipo de columna de nuevo al homodímero. Las alícuotas de elución fueron analizadas mediante PAGE-SDS, donde de nuevo se observa un gran grado de pureza de la proteína recombinante.

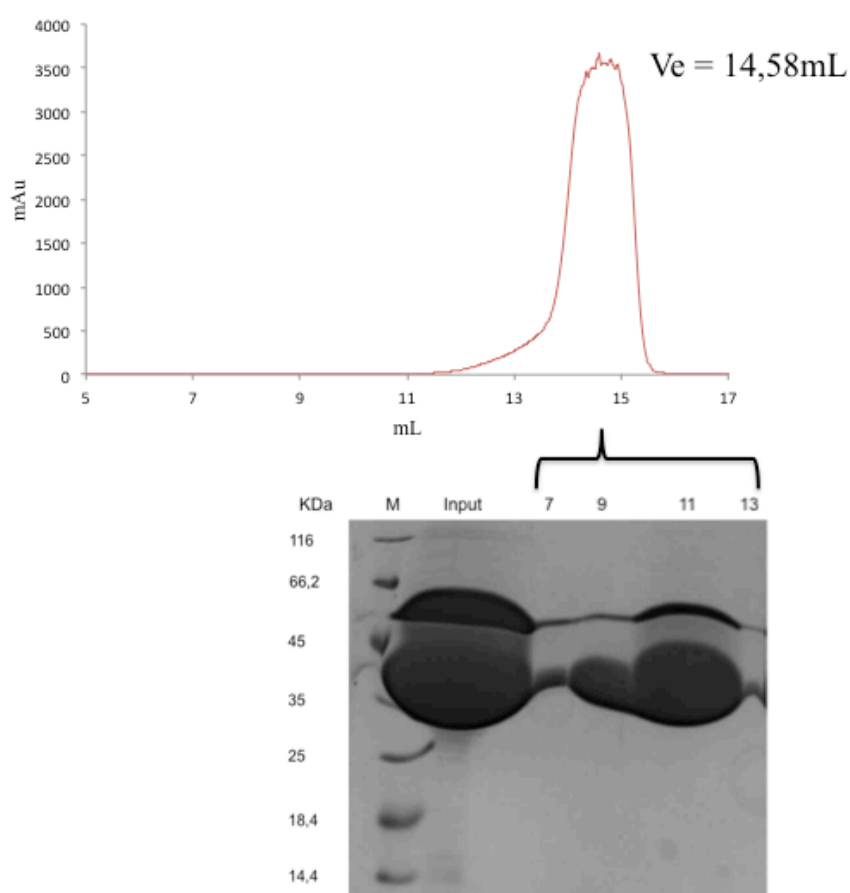


Figura 4.7. Purificación mediante filtración en gel de CsdA H120A. Se muestra el cromatograma generado, representando las miliunidades de absorbancia (mAU) a 280 nm frente al volumen. Se puede observar que el volumen de elución (V_e) es de 14,58 mL correspondiente al homodímero. En la PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En el gel de PAGE-SDS M es el marcador, *Input* es la muestra cargada en la columna y los números corresponden a las alícuotas analizadas.

Resultados

La purificación del mutante CsdA K250A se realizó del mismo modo que en el caso del anterior mutante. En la figura 4.8 se muestra el resultado de la IMAC. En este caso tan solo se muestra el análisis por PAGE-SDS, al llevarse el proceso de purificación empleando una bomba peristáltica.

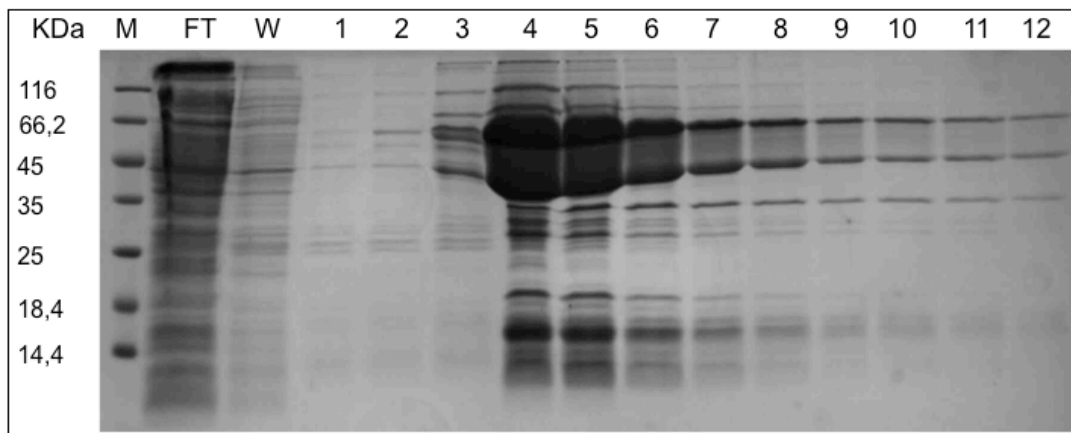


Figura 4.8. Purificación mediante IMAC del mutante CsdA K250A. Se muestran las alícuotas de la purificación en el gel de PAGE-SDS resultante, observándose que la mayoría de la proteína, sobre 45 kDa, corresponde a la recombinante CsdA K250A. M es el marcador, FT el *flow-through*, W lavado y los números corresponden a las alícuotas analizadas.

Por su parte en la figura 4.9 se muestra el resultado de la filtración en gel tras concentrar las alícuotas seleccionadas en el paso anterior. En este caso la proteína eluye a los 14,21 mL, dando de nuevo un homodímero como resultado. En el análisis por PAGE-SDS se puede observar la gran pureza de la proteína recombinante.

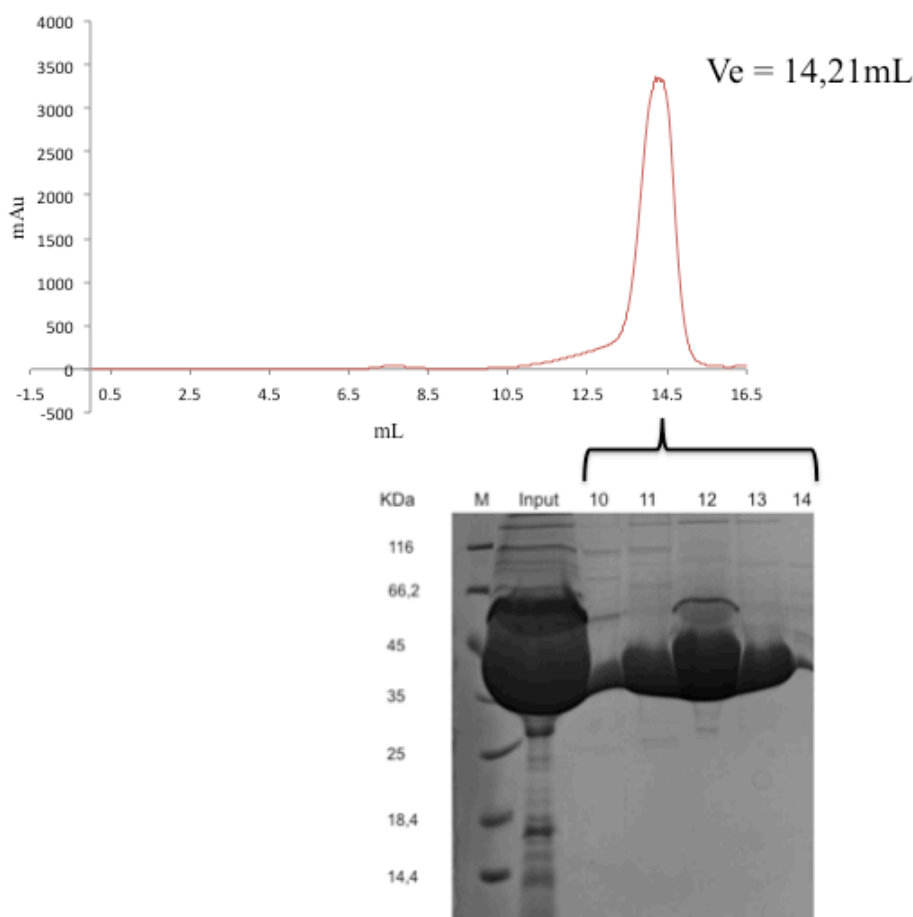


Figura 4.9. Purificación mediante filtración en gel del mutante CsdA K250A. Se muestran el cromatograma y la PAGE-SDS correspondiente al mutante CsdA K250A. En el cromatograma se representan las miliunidades de absorbancia (mAU) a 280 nm frente al volumen de elución. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M es el marcador, *Input* es la muestra cargada en la columna de filtración en gel y los números corresponden a las alícuotas analizadas.

4.1.6. Persulfuración de CsdA

Una vez purificada la proteína CsdA, 50 μ M de ésta fueron persulfurados mediante la incubación de la muestra de proteína con 5 mM de L-Cys (30 minutos 298 K). La proteína fue posteriormente purificada del exceso de L-Cys mediante filtración en gel en el tampón 10 mM Tris-HCl pH 8,0, 50 mM NaCl y 0,1 mM EDTA empleado la columna Superdex200 HR 16/60. La muestra obtenida en el pico de elución fue suplementada con 0,2 mM EDTA.

Resultados

Por otro lado se quiso saber si los mutantes generados de CsdA (H120A y K250A) eran capaces de captar los átomos de S de la L-Cys, como ocurre en el caso de la proteína nativa. Para ello se llevó a cabo el mismo proceso que en el caso de la proteína nativa, en cuanto a persulfuración, y se midieron las masas mediante MALDI-TOF. En la tabla 4.1 se muestran las masas obtenidas en cada caso.

Especie	Masa teórica (Da)	Masa experimental (Da)
CsdA	43.534,26	-
CsdA persulfurada	43.566,32	-
CsdA H120A	43.468,26	43.465,6
CsdA H120A persulfurada	43.500,33	43.499,5
CsdA K250A	43.477,26	43.475,0
CsdA K250A persulfurada	43.509,33	43.509,5

Tabla 4.1. Datos de masas para las especies sin modificar y modificadas (persulfuración) de CsdA. Se observa cómo en ningún de los casos la desulfurasa pierde la capacidad de tomar los átomos de S de la fuente.

En todos los casos se da la captura del átomo de S, habiendo en cada par (especie modificada y sin modificar), una diferencia aproximada de 32 Da, correspondiente a la masa molecular del átomo de azufre. Por su parte, las diferencias entre los mutantes y la proteína nativa se deben a las mutaciones, habiendo una reducción de la masa de 66 Da al mutar una His por una Ala y de 57 Da al mutar una Lys por una Ala.

4.1.7. Cristalización de CsdA nativa, persulfurada y oxidada

Los cristales de CsdA persulfurada fueron obtenidos mediante la técnica de difusión de vapor por gota sentada, empleando la solución de cristalización 0,2 M formato de magnesio, 18% (v/v) PEG 3.350 o 0,2 M tartrato disódico, 16% (v/v) PEG

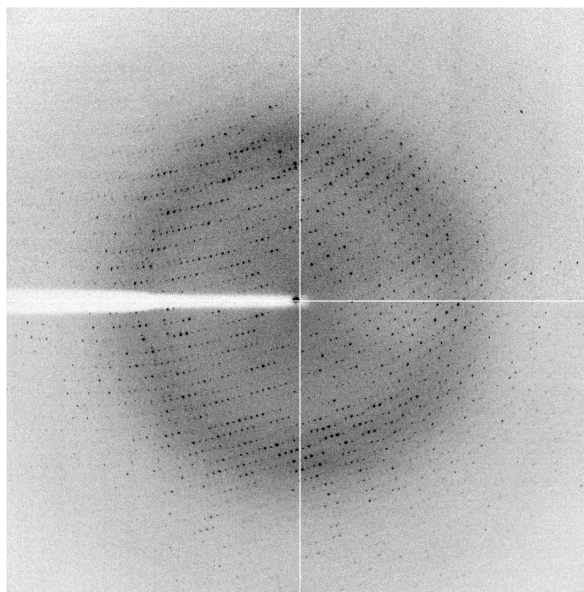
3.350. La concentración de proteína usada fue 20 mg/mL (previamente analizada por DLS). Los cristales se obtuvieron a 293 K, tardando en aparecer 1-4 días. Posteriormente fueron congelados del mismo modo que en el caso anterior.

Por último, los cristales de la proteína S-sulfatada fueron obtenidos a partir de los cristales de la proteína previamente persulfurada *in situ*, formándose después de 20 días, debido a la oxidación ambiental. De esta forma tanto en el caso de los cristales con proteína persulfurada, como con los cristales de los mutantes y la proteína S-sulfatada, éstos fueron complementados con 20% de glicerol para su congelación en nitrógeno líquido y transporte.

4.1.8. Colección y procesamiento de datos de difracción de rayos-X de la proteína nativa persulfurada y S-sulfatada

La colección de los datos fue llevada a cabo en las líneas ID23-2 (ESRF, Grenoble, Francia) para el caso de la proteína persulfurada, y en la línea ID14-1 (ESRF, Grenoble, Francia), en el caso de la proteína S-sulfatada. Ambos cristales, tanto el de la proteína en sus estado persulfurado como S-sulfatada, pertenecen al grupo espacial ortorrómbico $P2_12_12_1$, difractando ambos a una resolución entre 2,05-2,38 Å. En la figura 4.10 se muestra la imagen de difracción de los cristales de CsdA persulfurada y S-sulfatada. Todos los datos fueron procesados con XDS (Kabsch, 2010) y escalados con *Aimless* (Evans y Murshudov, 2013).

A



B

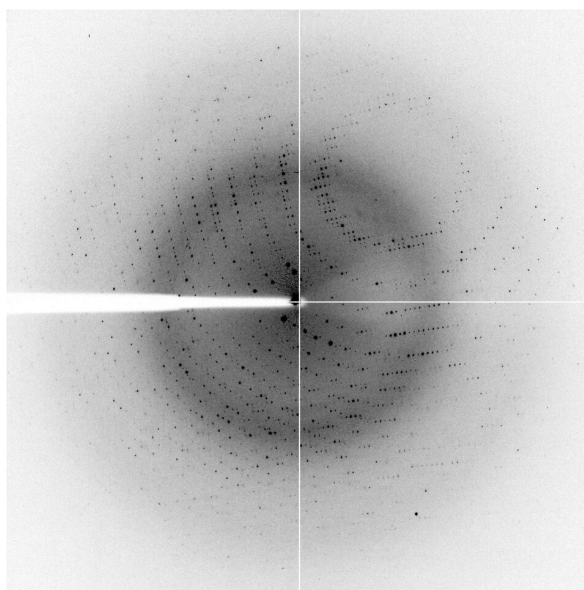


Figura 4.10. Imágenes de difracción de CsdA persulfurada y S-sulfatada. A, se muestra arriba la imagen de difracción obtenida para CsdA persulfurada. **B,** se muestra aabajo el cristal difractado correspondiente a CsdA S-sulfatada.

4.1.9. Resolución estructural y refinamiento de CsdA nativa persulfurada, oxidada

Las estructuras tanto de la proteína persulfurada como S-sulfatada fueron determinadas mediante MR, empleando como modelo la proteína nativa (pdb 5ft4) y usando el programa AMoRe (Trapani y Navaza, 2008). Los modelos generados de las dos estructuras fueron refinados mediante el empleo de *phenix.refine* (Afonine *et al.*, 2012), llevando a cabo una reconstrucción manual y una validación mediante *Coot* (Emsley *et al.*, 2010) y *MolProbity* (Chen *et al.*, 2010). El modelo final para la estructura persulfurada quedó a una resolución de 2,38 Å. Por su parte la estructura de la proteína S-sulfatada quedó a una resolución de 2,05 Å. En la tabla 4.2. se muestran las estadísticas obtenidas de la solución de las estructuras de CsdA persulfurada y S-sulfatada.

	(CsdAper) ₂	(CsdASSO ₃) ₂
PDB code	5ft4	5ft6
Data collection		
Wavelength (Å)	0.87260	0.93340
Resolution range (Å)	40.0-1.99 (2.07-1.99)	39.8-2.05 (2.11-2.05)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	64.09, 102.56, 71.25	75.97, 99.79, 105.75
β (°), α = γ = 90°	90, 114.45, 90	90, 90, 90
Total reflections	191,504 (17,351)	389,509 (19,170)
Unique reflections	56,541 (5,461)	50,628 (3,452)
Multiplicity	3.4 (3.3)	7.7 (3.4)
Completeness (%)	99.29 (96.14)	98.90 (92.50)
Mean I/σ (I)	11.18 (1.85)	19.6 (2.80)
Wilson B-factor	24.21	21.74
R-merge	0.105 (0.738)	0.101 (0.614)
R-meas^a	0.1243	0.1090
CC1/2^b	0.995 (0.585)	0.998 (0.815)
CC^c	0.999 (0.859)	1.000 (0.940)
Refinement		
R-work	0.1377 (0.2118)	0.1545 (0.2398)
R-free	0.1914 (0.2596)	0.1907 (0.2622)
# non-H atoms	6862	6988
# Protein	6190	6224
# Ligand atoms	88	68
# Water	584	694
Protein residues	798	799
RMS (bond) (Å)	0.012	0.003
RMS(angles)(°)	1.33	0.81
Ramachandran analysis		
Favored/Allowed/Outlier (%)	98.0/2.0/0.0	97.29/2.71/0.0
Clashscore	2.25	6.82
Average B-factor (Å²)	30.10	28.02
Protein	29.20	13.80
Ligands	49.90	41.50
Solvent	37.20	23.90

Tabla 4.2. Datos estadísticos de las estructuras de CsdA (estado persulfurado y S-sulfatado). (CsdAper)₂, dímero de CsdA persulfurado y (CsdASSO₃)₂, dímero S-sulfatado.

4.1.10. CsdA, la cisteína desulfurada de *E. coli*, presenta tres estados de oxidación de la cisteína catalítica (apo, persulfuración y S-sulfatada)

Como se ha comentado anteriormente, la proteína CsdA de *E. coli* pertenece a la familia I en cuanto el plegamiento de las enzimas dependientes de PLP (Mehta y Christen, 2000). Su actividad enzimática ha sido clasificada dentro de las cisteína desulfurasas, donde se encuentra en el grupo II (Mihara y Esaki, 2002), en base a la secuencia consenso alrededor del residuo catalítico Cys358 (tabla 1.4).

Al resolver las estructuras tridimensionales tanto del dímero de CsdA persulfurada como S-sulfatada, no se observan diferencias significativas en su estructura global, incluso con la forma apo, habiendo un r.m.s.d. de 0,5 Å (sobre 400 Cα) entre el dímero persulfurado y sin la modificación. En la figura 4.11 se muestra la estructura de la proteína persulfurada, siendo la misma para el caso de la proteína en su estado apo o S-sulfatado. Por lo tanto, la estructura general de CsdA se respeta, a pesar los cambios descritos en el centro activo, manteniéndose los dos dominios (grande y pequeño) que caracterizan a la proteína. Su estado oligomérico, homodímero, tampoco se ve alterado.

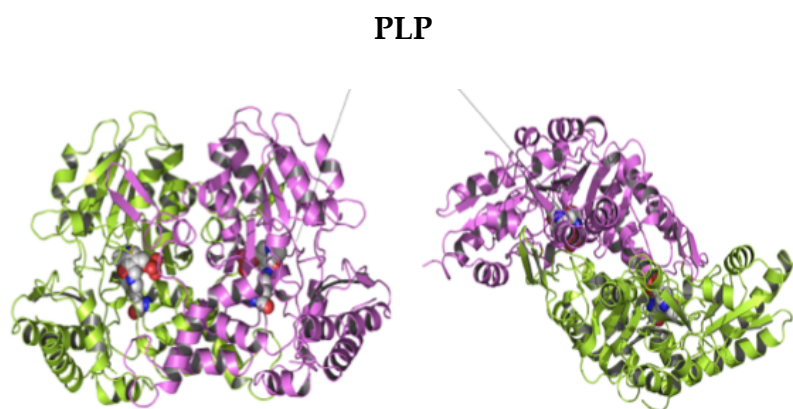


Figura 4.11. Estructura tridimensional del dímero de CsdA persulfurado (CsdAper)₂. Se muestra en dos orientaciones, 90° una respecto de la otra. Tras la incubación con cisteína, CsdA lleva a cabo la persulfuración a través del residuo catalítico Cys358. Este proceso no implica cambios significativos en la estructura, dándose pequeños ajustes en los aminoácidos alrededor del residuo catalítico. Se indica donde se encuentra la molécula del cofactor PLP en ambas orientaciones (Fernández *et al.*, 2016).

Resultados

En la estructura correspondiente a (CsdAper)₂ se puede observar los aminoácidos involucrados en el motivo donde se encuentra la Cys358. Están implicados los residuos entre 4-22 y entre 294-401 de la secuencia de la proteína, ambos en el dominio pequeño de ésta, entre los que se encontraría la secuencia consenso del grupo II de este tipo de proteínas (³⁵³RAGQHCA³⁵⁹), marcando la cisteína catalítica en negrita. Este dominio, aunque cumple la misma función en cada una de las cisteína desulfurasas, (permitir la captura y la transferencia de átomos de azufre) difiere en su grado de flexibilidad, presentando en el caso de IscS una alta movilidad (Shi *et al.*, 2010), hecho no observado en las otras dos cisteína desulfurasas descritas en *E. coli* (CsdA y SufS), donde este motivo se presenta rígido. De hecho en la estructura del complejo (CsdAapo-CsdE)₂ no se observa cambio estructural en el motivo (Kim y Park, 2013).

En cuanto a la estabilización del motivo persulfuro en la proteína (CsdAper)₂, la Cys358 se ve apoyada por la formación de puentes de hidrógeno provenientes del grupo imidazol de la His119 y de la amida de la cadena principal de la Lys250* (de la otra subunidad del homodímero). Las interacciones sobre el Sδ de la Cys358 modificada, hacen que se oriente de forma que favorece la reacción de transpersulfuración hacia la molécula de CsdE (figura 4.12).

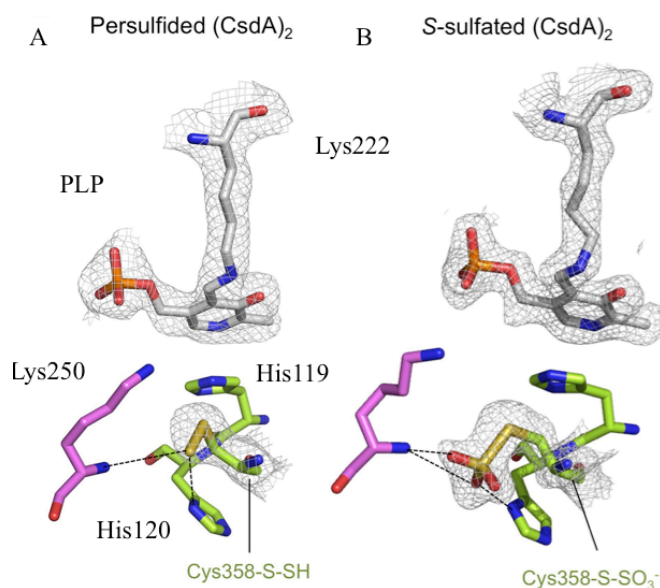


Figura 4.12. Estructura del centro activo de CsdA en diferentes estados de oxidación. A Dímero (CsdAper)₂ (persulfurado) y B dímero (CsdAss)₂ (S-sulfatado). Los aminoácidos se muestran en colores CPK, excepto los carbonos. Éstos se muestran en gris en el caso de Lys222-PLP y en verde en el caso de CsdA. Los mapas de densidad electrónica a 1 σ se muestran alrededor de la Lys222-PLP y de los residuos de cisteína catalíticos (Cys358 en el caso de CsdA) (Fernández *et al.*, 2016).

La última de las estructuras obtenidas del dímero de CsdA, la oxidación del dímero de CsdA persulfurado, se originó espontáneamente ante la exposición con el aire, dando como resultado una especie S-sulfatada (CsdAss)₂ (figura 4.12), la cual no presenta cambios estructurales significativos con la forma persulfurada al darse un r.m.s.d. de 0,436 Å (sobre 796 Cα). Sin embargo, se ha observado que esta forma de la proteína no puede ser revertida por la maquinaria celular de la bacteria, impidiendo su corrección (Reddie y Carroll, 2008). No obstante, los residuos que interaccionan con la Cys358 S-sulfatada son los mismos que en el caso de (CsdAper)₂, acomodando la cadena lateral de la cisteína oxidada sobre los mismos residuos del centro activo que estabilizan la forma funcional. De esta forma se entiende, como bajo circunstancias de oxidación, CsdA es incapaz de ceder átomos de azufre tanto a CsdE como a SufE, siendo en este caso el sistema Suf el descrito como capaz de actuar bajo este tipo de condiciones (Ayala-Castro *et al.*, 2008 ; Dai y Outten, 2012), dejando sistemas como CSD bloqueados bajo esas condiciones celulares.

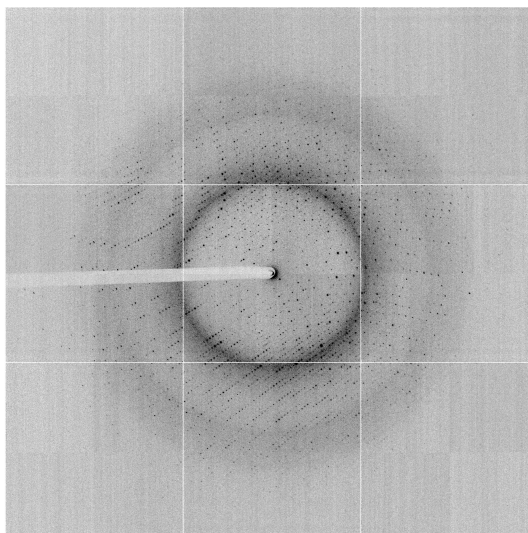
4.1.11. Cristalización de los mutantes CsdA H120A y CsdA K250A

Los cristales de los mutantes de CsdA (H120A y K250A) fueron obtenidos de igual manera que los cristales de CsdA nativa, mediante la técnica de difusión de vapor por gota sentada y empleando las mismas soluciones. La concentración de proteína usada fue 20 mg/mL (previamente analizada por DLS). Los cristales tardaron en aparecer 1-4 días y se empleo glicerol como crio-protector.

4.1.12. Colección y procesamiento de datos de difracción de rayos-X de los mutantes de CsdA

En el caso de los mutantes de CsdA la colección de los datos fue llevada a cabo en la línea PROXIMA 2A (Soleil, París, Francia). Los cristales de ambos mutantes pertenecen al grupo espacial ortorrómbico P2₁2₁2₁, difractando el cristal del mutante H120A a 1,74 Å y el del mutante K250A a 1,25 Å. En la figura 4.13 se muestra la imagen de difracción de los cristales los dos mutantes de CsdA. Todos los datos fueron procesados con XDS (Kabsch, 2010) y escalados con *Aimless* (Evans y Murshudov, 2013).

A



B

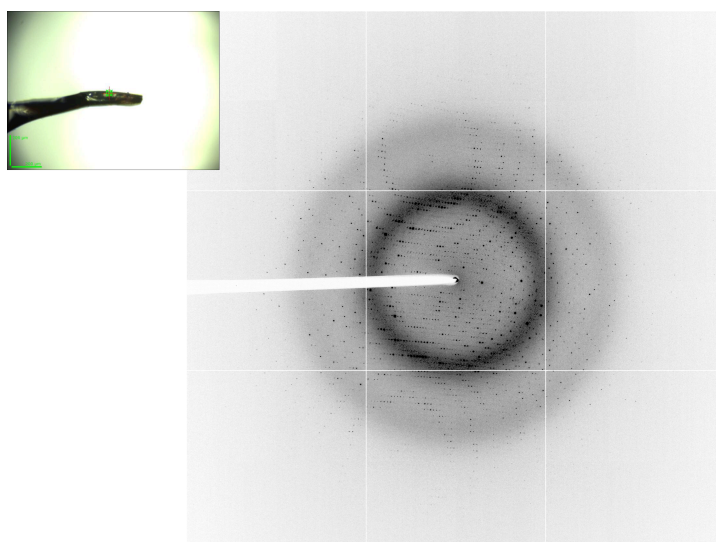


Figura 4.13. Imágenes de difracción de los mutantes de CsdA H120A y K250A. Se muestra arriba (A) la imagen de difracción del mutante H120A y abajo (B) la del mutante K250A. En este caso solo se muestra la imagen del cristal del mutantes CsdA K250A.

4.1.13. Resolución estructural y refinamiento de los mutantes de CsdA

En cuanto a las estructuras de los mutantes de CsdA, éstos fueron resueltos a una resolución de 1,74 Å en el caso del mutante H120A, y de 1,25 Å en el caso del mutante

K250A, del mismo modo que la proteína nativa. En la tabla 4.3 se muestran las estadísticas cristalográficas de ambos.

	CsdA H120A	CsdA K250A
PDB code	-	-
Data collection		
Wavelength (Å)	0.9801	0.9801
Resolution range (Å)	56.51-1.74 (1.8-1.74)	38.02-1.248 (1.29-1.25)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	64.183, 102.933, 71.787	64.04, 102.84, 71.42
β (°), $\alpha = \gamma = 90^\circ$	90, 113.622, 90	90, 114.13, 90
Total reflections	413,281 (41,088)	2,541,564 (99,347)
Unique reflections	87,737 (8,714)	227,346 (21,031)
Multiplicity	4.7 (4.7)	11.2 (4.7)
Completeness (%)	100 (100)	97 (90)
Mean I/σ (I)	10.69 (2.21)	13.15 (1.51)
Wilson B-factor	12.54	12.82
R-merge	0.12 (0.732)	0.09819 (1.137)
R-meas^a	0.1353 (0.8246)	0.1026 (1.28)
CC1/2^b	0.995 (0.691)	0.999 (0.501)
CC^{*c}	0.999 (0.904)	1.000 (0.817)
Refinement		
R-work	0.1613 (0.2415)	0.1234 (0.2697)
R-free	0.1868 (0.2600)	0.1499 (0.3083)
# non-H atoms	7581	7413
# Protein	6314	6444
# Ligand atoms	30	64
# Water	1237	905
Protein residues	803	805
RMS (bond) (Å)	0.006	0.012
RMS(angles)(°)	0.98	1.41
Ramachandran analysis		
Favored/Allowed/Outlier (%)	97.0/2.5/0.0	97/2.6/0.0
Clashscore	6.73	3.70
Average B-factor (Å²)	18.50	18.75
Protein	16.07	16.87
Ligands	37.10	42.76
Solvent	30.49	30.48

Tabla 4.3. Datos estadísticos de la solución de las estructuras de los mutantes de CsdA H120A y K250A.

4.1.14. Los mutantes de CsdA estudiados (H120A y K250A) no muestran cambios significativos en la estructura respecto a la proteína nativa

No se observan disfunciones en la capacidad de los mutantes para la toma de azufre del aminoácido fuente, L-Cys, por lo que éstos son capaces de autopersulfurarse. Estos residuos (H120A y K250A) no parecen fundamentales en el proceso de toma de azufre del aminoácido L-Cys. En las figuras 4.14 y 4.15 se representan las estructuras cristalográficas de ambos mutantes (H120A y K250A, respectivamente), donde no se observan diferencias significativas en cuanto al r.m.s.d, respecto de la proteína nativa persulfurada y sin persulfurar, por lo que estas mutaciones no influyen en la estructura tridimensional de la proteínas.

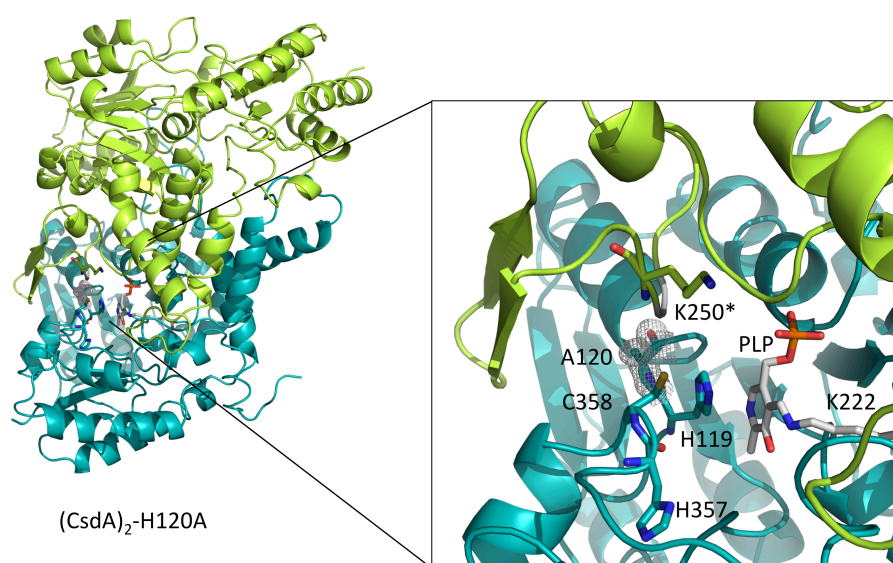


Figura 4.14. Estructura cristalográfica del dímero (CsdA)₂-H120A. La estructura se muestra a una resolución de 1,74 Å. Se puede observar el cofactor (PLP) y los residuos involucrados en el centro activo en forma de barras. La cadena lateral del residuo mutado, Ala120, se puede apreciar en el mapa de densidad electrónica 2Fo-Fc a un nivel de 1 σ .

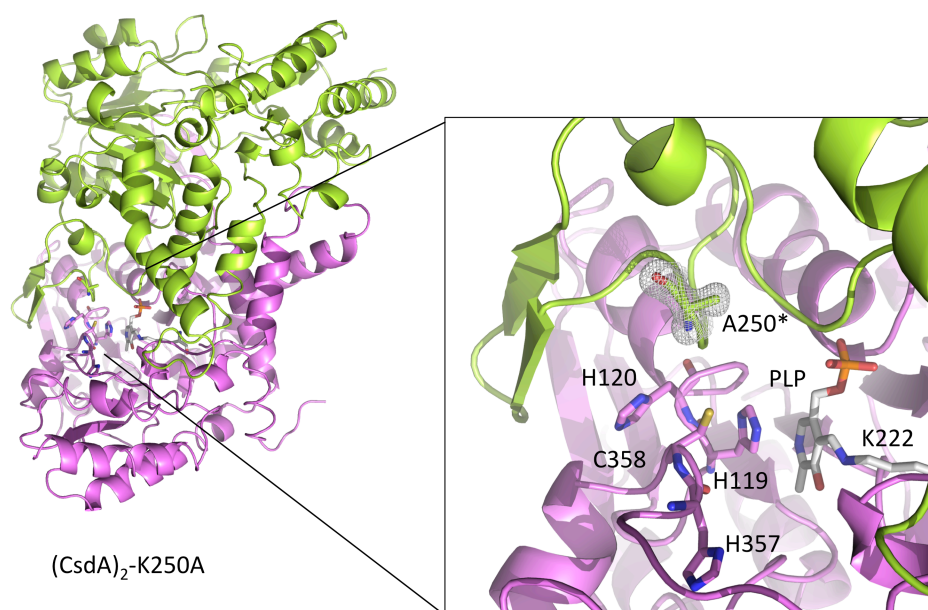


Figura 4.15. Estructura cristalográfica del dímero (CsdA)₂-K250A. La estructura se muestra a una resolución de 1,25 Å. Se puede observar el cofactor en su forma PLP y los residuos involucrados en el centro activo en forma de barras. La cadena lateral del residuo mutado, Ala250, se puede apreciar en el mapa de densidad electrónica 2Fo-Fc a un nivel de 1 σ .

En cuanto al cofactor de este tipo de enzimas, el PLP, éste forma una base de *Schiff* con el residuo catalítico Lys222 en cada una de las subunidades de CsdA, formando una aldimina interna, participando de forma directa tanto en el proceso de desulfuración como en el de transpersulfuración (figuras 4.14 y 4.15).

4.2. CsdE

4.2.1. Clonaje de CsdE

El gen codificante para la CsdE se obtuvo del genoma de *E. coli* (XL1blue) mediante la técnica PCR y el uso de los oligonucleótidos específicos para este gen (anexo I, tabla 2). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes se clonó en el vector de expresión con resistencia a ampicilina pET21b. En este caso tanto el vector como los insertos fueron digeridos con las enzimas *NdeI*/*XhoI*. Una vez obtenida la construcción pET21b-CsdE, se transformó en la cepa de clonaje de XL1blue, creciendo colonias en placas de LB con agar al 1,5% con ampicilina 100 µg/mL. Al comprobar que la construcción era correcta mediante

Resultados

PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen (anexo I, tabla 2), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.2.2. Expresión de CsdE

Posteriormente se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Ésta se realizaron en 1 mL de LB más ampicilina a la concentración mencionada anteriormente. De esta forma las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación se mantiene la temperatura y se induce el cultivo con la adición de 0,5 mM IPTG (3 horas 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonificado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 µL del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más ampicilina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin embargo, el tiempo de expresión se llevó a 5 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

La expresión de la proteína doblemente marcada (¹³C, ¹⁵N) se realizó del mismo modo cambiando simplemente el medio de expresión. Para ello se empleó el medio mínimo de cultivo M9 (100 mM Na₂HPO₄, 20 mM KH₂PO₄, 10 mM NaCl, 25 mM D-glucosa (¹³C), 20 mM NH₄Cl (¹⁵N), 1 mM MgCl₂, 0,1 mM CaCl₂ y 35 µM tiamina). El resto de los parámetros se respetaron.

4.2.3. Purificación de CsdE

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular generado fue resuspendido en el tampón de lisis (20 mM Tris-HCl pH 7,5, 150 mM NaCl, 25 mM imidazol y 2 mM β-mercaptoetanol), al que se le añade 10 mM de PMSF. Una vez resuspendido el sedimento celular, éste es sonificado (2,5 minutos 277 K) en ciclos de 30 segundos de

sonicación y 20 segundos de reposo. A continuación el lisado es centrifugado a 10.000 g (20 minutos, 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 μm de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de un ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando gradualmente la concentración de imidazol hasta 250 mM, sin cambiar ninguno de los otros componentes del tampón. La proteína eluyó a 125 mM de imidazol (figura 4.16). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 15 kDa, correspondiente a la masa molecular del monómero de CsdE (figura 4.16).

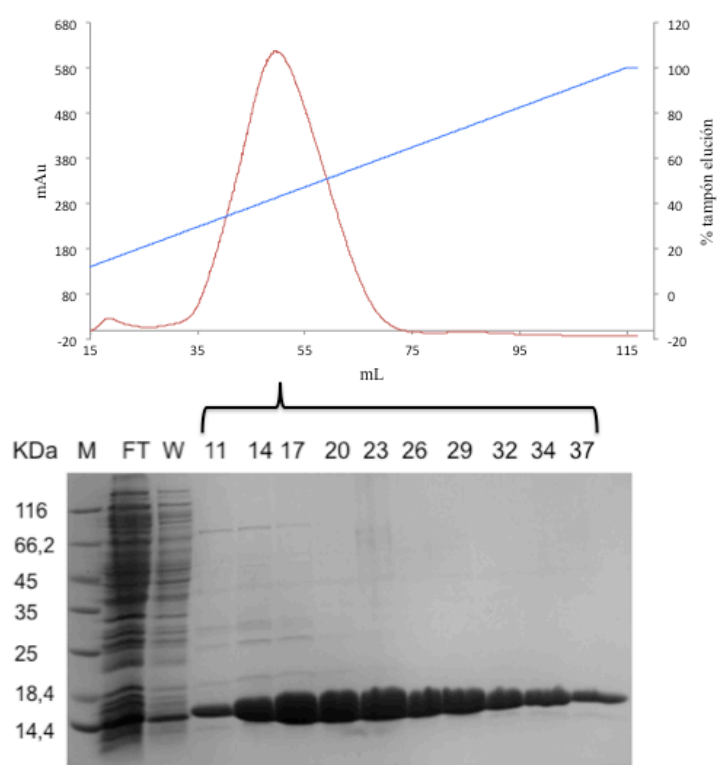


Figura 4.16. Purificación mediante IMAC de CsdE. En el cromatograma se representan las miliunidades de absorbancia (mAU) a 280 nm (línea roja) frente al volumen de elución. La línea azul representa la concentración del tampón de elución, observándose el gradiente. En la PAGE-SDS se muestran las muestras de la purificación, siendo la mayoría de la proteína, sobre 15 kDa, correspondiente a la recombinante CsdE. Se puede observar la gran pureza de la muestra. En el gel de PAGE-SDS M es el marcador, FT el *flow-through*, W el lavado y los números corresponde a las alícuotas analizadas.

Resultados

A continuación la proteína es concentrada mediante el uso de concentradores (*Millipore*). Tras concentrar la proteína, ésta es cargada en una columna de filtración en gel (*HiLoad 16/60 Superdex75, GE Healthcare*), equilibrada previamente con el tampón 20 mM Tris-HCl pH 7,5, 150 mM NaCl y 2 mM β -mercaptoetanol. La filtración en gel fue llevada a cabo a 1 mL/min, eluyendo la proteína al volumen de elución esperado para el monómero, 108,92 mL (figura 4.17), de aproximadamente 15 kDa. De nuevo las alícuotas del pico de elución obtenido fueron analizadas mediante PAGE-SDS, donde se puede apreciar el grado de pureza de la proteína.

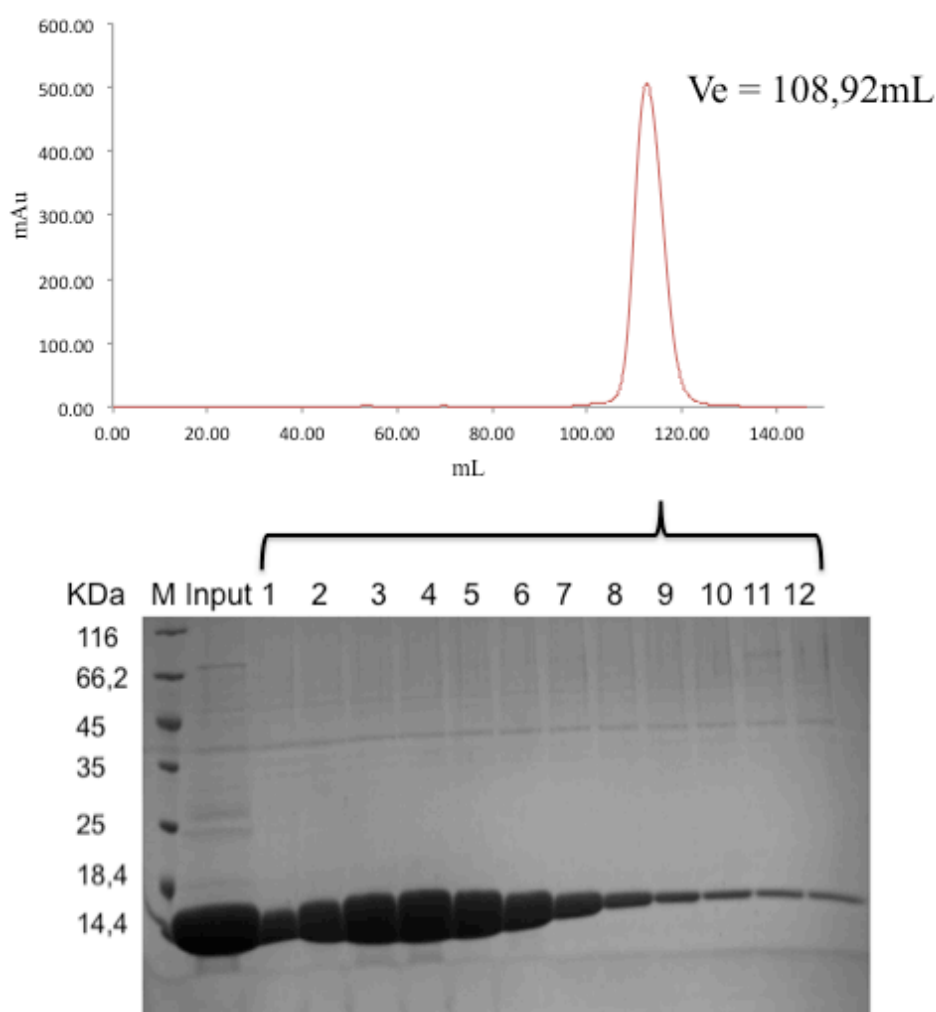


Figura 4.17. Purificación mediante filtración en gel de CsdE. Se puede observar que el volumen de elución (V_e) es de 108,92 mL correspondiente al monómero de la proteína. En esta ocasión se representa de nuevo en rojo las miliunidades de absorbancia (mAU) a 280 nm frente al volumen. En la PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M es el marcador, *Input* la muestra cargada en la columna y los números son las alícuotas analizadas.

4.2.4. Clonaje de mutantes de CsdE

El clonaje de los diferentes mutantes generados de CsdE fue llevado a cabo por Juan Quintana (CIB, CSIC), empleando el método de empalme-extensión. El molde usado en la reacción de PCR para generar los insertos con la mutación introducida fue pET21b-CsdE. Los oligonucleótidos empleados se muestran en la tabla 2 del anexo I.

Una vez obtenidos los fragmentos de PCR y digeridos correctamente con las enzimas de restricción correspondientes, en todos los casos se emplearon las mismas enzimas, se clonó en el vector de expresión con resistencia a ampicilina pET21b. En este caso tanto el vector como los insertos fueron digeridos con las enzimas *NdeI*/*XhoI*. Una vez obtenida la construcción, se transformó en la cepa de clonaje de *XL1blue*, creciendo colonias en placas de LB con agar al 1,5% con ampicilina 100 µg/mL. Una vez comprobado que la construcción era correcta mediante PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen (anexo I, tablas 1 y 2), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.2.5. Expresión y purificación de mutantes de CsdE

Estos dos pasos se realizaron del mismo modo que en la proteína nativa, obteniendo los mismos resultados en cuanto a cantidad y calidad de proteína recombinante y estado oligomérico, además de no observarse en ninguno de los casos pérdida de estabilidad de la proteínas (agregados o precipitados). Los mutantes expresados y purificados de CsdE fueron Arg21Glu/Asn22Glu (R21E/N22E), Arg86Glu (R86E), Arg129Asp (R129E), Tyr34Glu/Arg35Glu (Y34E/R35E), Lys42Glu/Arg86Glu (K42E/R86E), Tyr34Glu/Arg35Glu/Arg129Glu (Y34E/R35E/R129E) y Lys42Glu/Arg86Glu/Arg129Glu (K42E/R86E/R129E).

Se muestra una purificación del mutante CsdE R21E/N22E, representativa del resto de mutantes de CsdE. De esta forma tras el procesamiento del sedimento celular generado, el primer paso de purificación fue una IMAC, con la que se obtuvo un cromatograma, lo que permitió el análisis de las muestras eluidas mediante PAGE-SDS. En la figura 4.18 se muestra el resultado de este paso.

Resultados

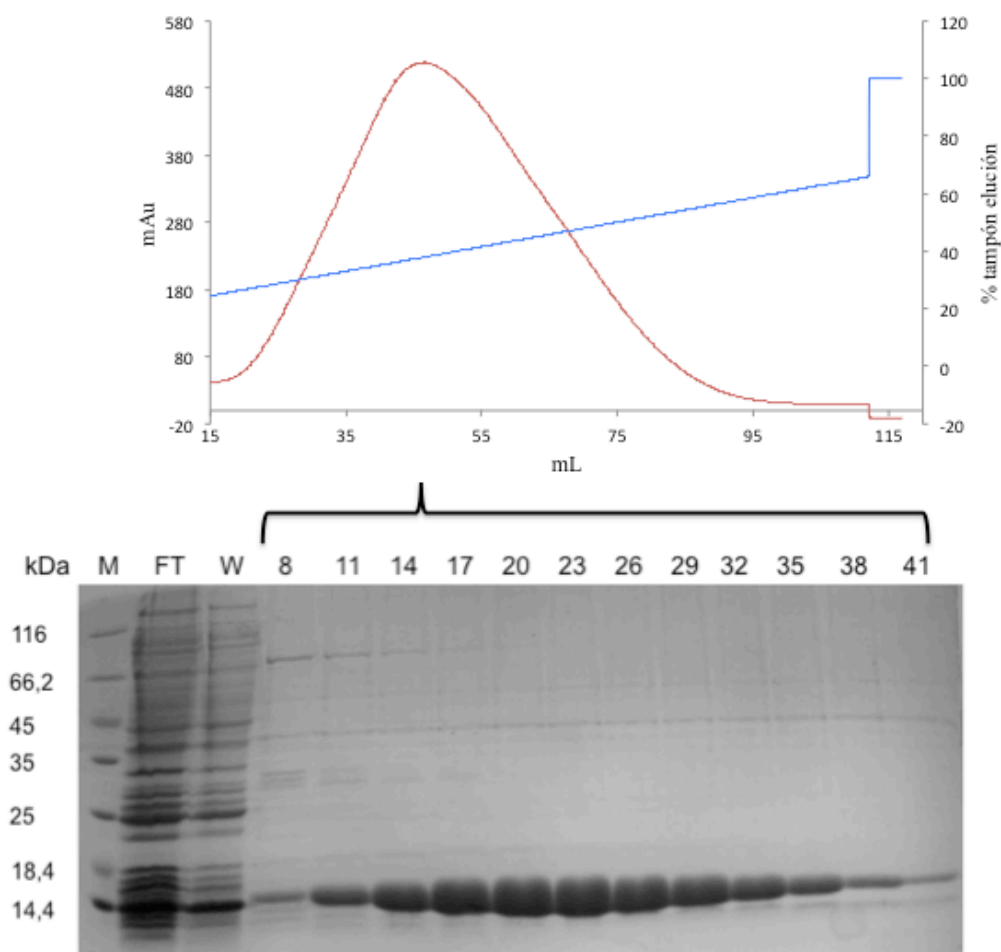


Figura 4.18. Purificación mediante IMAC del mutante R21E/N22E de CsdE. En el cromatograma se representan las miliunidades de absorbancia (mAU) a 280 nm (línea roja) frente al volumen de elución. La línea azul representa la concentración del tampón de elución, observándose el gradiente. En la PAGE-SDS se muestran las muestras de la purificación, siendo la mayoría de la proteína, sobre 15 kDa, correspondiente a la recombinante CsdE (R21E/N22E). Se puede observar la gran pureza de la muestra. En el gel de PAGE-SDS M es el marcador, FT el *flow-through*, W el lavado y los números corresponde a las alícuotas analizadas.

A continuación la proteína seleccionada de la elución es concentrada y tratada del mismo modo que en la proteína nativa con el fin de llevar a cabo una filtración en gel bajo las mismas condiciones que en el caso de la proteína nativa. Se obtuvo un volumen de elución de 112,08 mL en la misma columna que en el caso de la proteína nativa, comportándose de nuevo la proteína como un monómero. En la figura 4.19 se muestra el resultado de este último paso de purificación, donde se observa que la mayoría de la proteína se encuentra sobre 15 kDa, en un alto grado de pureza.

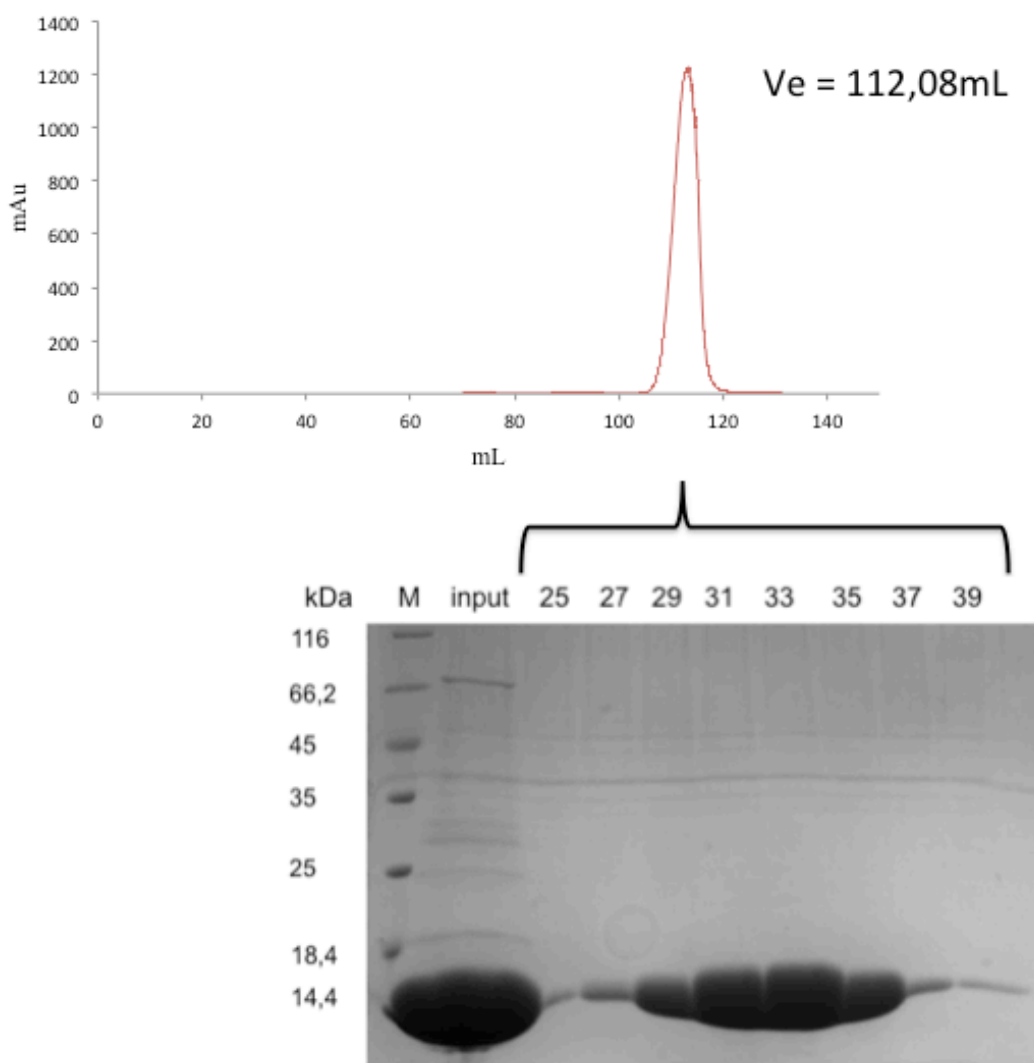


Figura 4.19. Purificación mediante filtración en gel del mutante R21E/N22E CsdE. Se puede observar que el volumen de elución (V_e) es de 112,08 mL correspondiente al monómero de la proteína. En esta ocasión se representa de nuevo en rojo las miliunidades de absorbancia (mAU) a 280 nm frente al volumen. En la PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M es el marcador, *Input* la muestra cargada en la columna y los números son las alícuotas analizadas.

En la figura 4.20 se muestra un gel de PAGE-SDS, donde se muestra una alícuota del último paso de purificación de cada uno de los mutantes de CsdE ensayados en esta tesis con el fin de mostrar su grado de pureza.

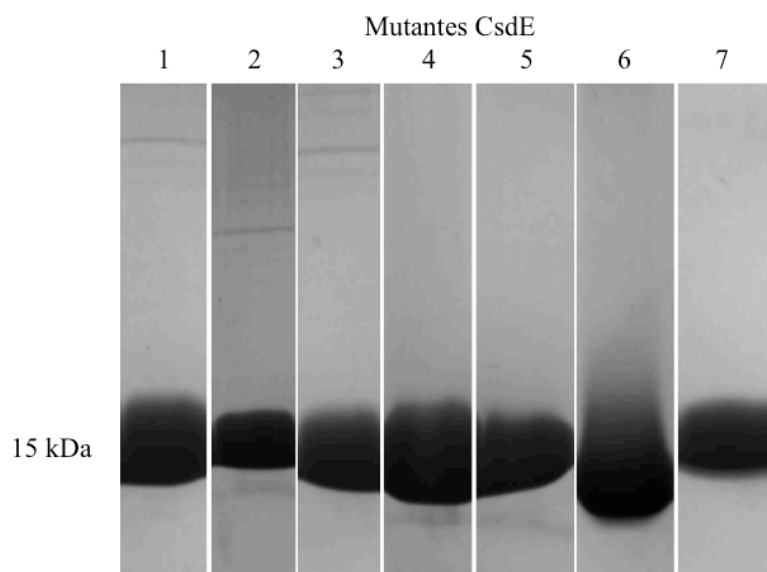


Figura 4.20. Análisis por PAGE-SDS de una muestra del último paso de purificación de cada uno de los mutantes de CsdE. Se observa la gran pureza de cada uno de los mutantes purificados. Los números indican cada uno de los mutantes ensayados: 1 (R21E/N22E), 2 (R86E), 3 (R129E), 4 (Y34E/R35E), 5 (K42E/R86E), 6 (Y34E/R35E/R129E) y 7 (K42E/R86E/R129E). Todos ellos a la altura del peso molecular esperado, 15 kDa.

4.3. Complejo de CsdA-CsdE

4.3.1. Ensamblado y purificación de CsdA-CsdE

El ensamblaje del complejo (CsdA-CsdE)₂ se llevó a cabo a partir de las proteínas previamente purificadas. Para ello fueron mezcladas en un ratio 1:1,5 (CsdA:CsdE) y purificadas mediante filtración en gel en la columna Superdex200 10/300GL previamente equilibrada con el tampón 10 mM HEPES pH 7,4, 50 mM NaCl y 0,1 mM EDTA. La purificación fue llevada a cabo gracias a la conexión de la columna a un ÄKTA Prime, realizando la purificación a 0,5 ml/min.

4.3.2. Persulfuración del complejo CsdA-CsdE

Una vez formado el complejo, la proteína obtenida del pico de elución del punto anterior fue persulfurada mediante la adición de 5 mM de L-Cys a la muestra, incubándola durante 30 minutos a 298 K. El exceso de CsdE fue eliminado de la muestra gracias a la purificación mediante una filtración en gel, llevada a cabo del

mismo modo que en el apartado anterior, donde el complejo fue fácilmente separado del monómero de CsdE sobrante.

4.3.3. Caracterización de la interacción CsdA-CsdE mediante SPR

Con el fin de caracterizar la interacción entre estas dos proteínas se llevaron a cabo experimentos de SPR (Resonancia de Plasmón de Superficie), mediante el uso de un BIAcore X100 (GE *Healthcare*) empleando el kit de captura de anticuerpos-His (GE *Healthcare*). El anticuerpo anti-His se inmovilizó en un chip CM5 empleando una amina acoplada químicamente tanto en la referencia como en los canales de las muestras, para un nivel de 13.000-15.000 RU (unidades de respuesta). La proteína CsdE-His₆ purificada (ya sea la nativa o los diferentes mutantes) fueron capturados al comienzo de cada experimento entre 60-120 RU, aplicando posteriormente (CsdA)₂ sobre el chip, siendo en este caso CsdE el ligando y CsdA el analito. La respuesta máxima teórica (R_{\max}) se estimó de acuerdo con el ratio de masa molecular entre CsdE (16,7 kDa) y CsdA (40 kDa), i.e. $100 \times 16,7/40 = 41,75$ (por cada 100 RU de CsdE-His₆ capturada). Los datos fueron analizados con el programa BIAevaluation en los modos de afinidad y cinéticas.

En cuanto a las cantidades de muestra empleadas, en el caso de CsdE o de sus variantes fue de 100 μ M, mientras que la de CsdA fue de 8-5.000 nM en el caso de las interacciones con CsdE nativa y CsdE_R21E/N22E, y de 0,16-100 μ M en el caso de CsdE_R86E, CsdE_R129E, CsdE_Y34E/R35E, CsdE_K42E/R86E, CsdE_Y34E/R35E/R129E y CsdE_K42E/R86E/R129E.

Esta técnica permite detectar la cantidad de analito ((CsdA)₂) que se une al elemento fijado en el chip, el ligando (CsdE), de forma que la muestra (CsdA)₂ fluye a través de unos canales interaccionando con CsdE, midiendo el cambio de intensidad del haz de luz reflejada sobre el chip al modificar el ángulo de incidencia. En la figura 4.21 y en la tabla 4.4 se muestran los resultados de este experimento, así como las K_d obtenidas en cada caso.

Resultados

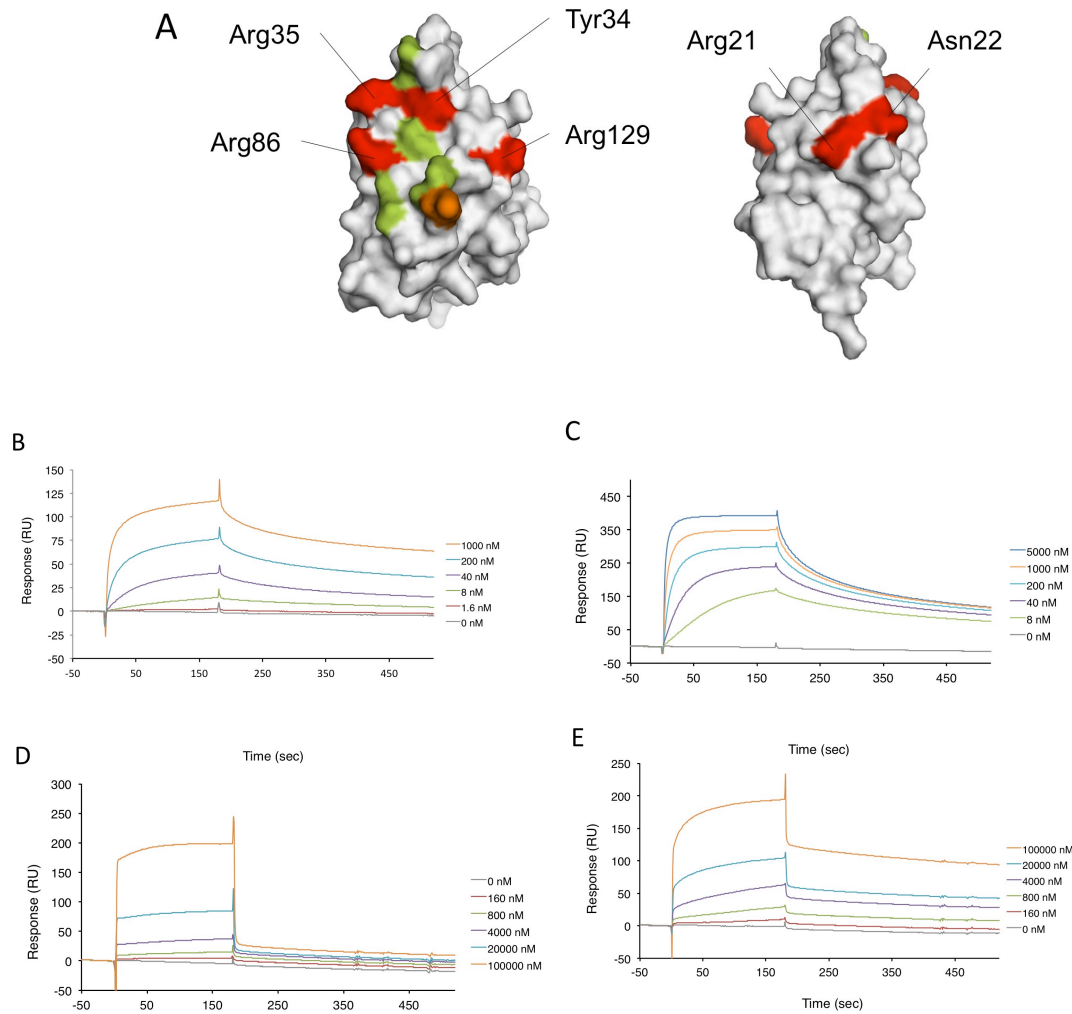


Figura 4.21. Sensorgrama de SPR de la interacción entre CsdE-His₆ nativa y mutantes con CsdA. **a** Mapa de superficie de los residuos de CsdE que interaccionan con CsdA (en rojo y verde). En rojo se muestran los residuos que fueron mutados para los ensayos de unión por SPR, incluyendo los residuos críticos para esta unión (izquierda) como aquellos empleados como control incorporados en la cara opuesta de CsdE, lejos de la zona de interacción con CsdA (derecha). La Cys61 de CsdE se muestra en naranja. **b-d** Diagramas de sensibilidad de la valoración con CsdA y CsdE inmovilizado sobre la superficie. **b** CsdE nativa, **c** CsdE_R21E/N22E (controles mutantes de carga inversa), **d** CsdE_R86E y **e** CsdE_R129E. Los ensayos de valoración con los dobles y triples mutantes de CsdE (CsdE_Y34E/R35E, CsdE_K42E/R86E, CsdE_Y34E/R35E/R129E y CsdE_K42E/R86E/R129E) permiten la abolición de unión de CsdA por completo. Las RU (unidades de respuesta) de la unión de CsdA se representan con el tiempo, con incrementos de la concentración de CsdA mostrados como variación del color.

CsdE ^a	K _d (nM)
nativa	125
R21E/N22E	177
R86E	52.000
R129E	29.940
DM/TM ^b	nd

Tabla 4.4. Cálculo de la constante de disociación (K_d) para (CsdA-CsdE)₂ medido por SPR. a Variante de CsdE empleada para el experimento. La K_d fue calculada gracias a la respuesta en equilibrio con diferentes concentraciones de analito (CsdA). **b** Dobles (DM) y triples mutantes (TM) de CsdE, gracias a los cuales se rompió totalmente la interacción entre ambas proteínas.

4.3.4. Cambio conformacional entre CsdE libre (RMN, pdb 1ni7) y CsdE unida a CsdA

Para estudiar el cambio conformacional que CsdE experimenta en su interacción con CsdA realizamos un estudio ¹H-¹⁵N HSQC de RMN con CsdE doblemente marcada (¹⁵N y ¹³C) en presencia de CsdA. Para el experimento se empleó una solución de CsdE doblemente marcada a una concentración de 90 μM, a la que se fue añadiendo una disolución concentrada de CsdA sin marcar (en exceso molar 4 veces mayor a la concentración de CsdE). Facilitó mucho el experimento el pequeño tamaño de CsdE (15 kDa). El experimento de (¹H, ¹⁵N) HSCQ permite seguir la interacción midiendo la intensidad del pico integrado para todos los grupos NH de CsdE por residuo afectado al añadir CsdA a la mezcla, permitiendo de este modo monitorizar la interacción de CsdE y CsdA durante la formación del complejo (CsdA-CsdE)₂. A pesar de disponer de la estructura de CsdE previamente resuelta por RMN (bmr5630, PDB 1ni7), el espectro de la CsdE monomérica fue reasignado de nuevo para compensar las diferencias con la asignación previa (Liu *et al.*, 2005), ya que éstas podían dificultar el seguimiento de los picos del espectro obtenido de la CsdE libre y formando complejo heterotetramérico (CsdA-CsdE)₂ en el presente experimento. En la figura 4.22 se muestra la

Resultados

representación de la estructura de CsdE por RMN, así como los residuos implicados en la interacción con CsdA.

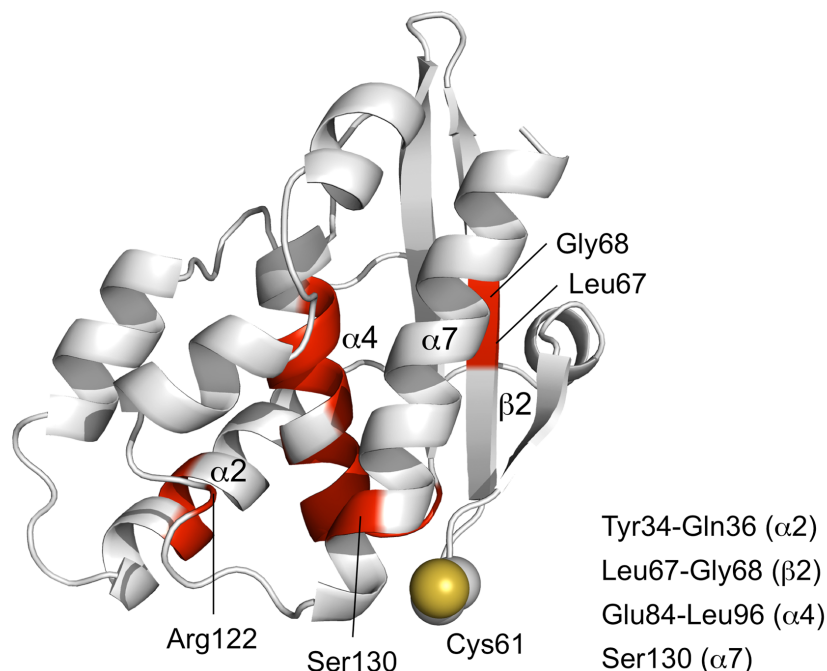


Figura 4.22. Representación de la estructura de CsdE por RMN (pdb 1ni7). Los aminoácidos cuya señal se ve disminuida en el espectro de ^1H - ^{15}N HSQC tras la adición de la proteína CsdA se marcan en rojo. La Cys61 está indicada como una esfera (amarillo para el átomo de S y blanco para el de C) (Fernández *et al.*, 2016).

Los residuos de CsdE afectados al formar el complejo mencionado son Tyr34, Arg35, Gln36, Glu84, Val88, Arg89, Gly90, Leu92, Ala93, Val94 y Leu96. Estos residuos mapean en las 3 α -hélices de CsdE que interactúan con CsdA al formar el complejo encajando con lo observado por cristalografía.

4.3.5. SAXS del complejo CsdA-CsdE

Los experimentos de SAXS se llevaron a cabo en la línea bioSAXS BM29 (ESRF, Grenoble, Francia). Para ello se analizaron CsdA₂, CsdE y (CsdA-CsdE)₂, todos ellos previamente purificados y analizados mediante DLS. De esta forma 30 μ L de proteína, entre las concentraciones 0,5-5 mg/mL fueron analizados para cada uno de los tres casos señalados. Las muestras se expusieron a rayos-X a 298 K, colectando los datos de dispersión. Se tomaron también medidas del tampón en el que se encontraban las proteínas, tanto antes como después de las medidas con el fin de eliminar el fondo. La dispersión, proporcional a la intensidad a ángulo 0 $[I(0)]$ y el radio de giro (Rg) se evaluaron mediante la aplicación de la ley de Guinier. El volumen de partícula de hidratación fue calculado mediante la ley de Porod y el máximo tamaño de partícula se determinó mediante GNOM (Svergun, 1992) usando PRIMUS (Konarev *et al.*, 2003).

La baja resolución de los modelos de CsdA₂, CsdE y (CsdA-CsdE)₂ fueron calculados mediante el uso de DAMMIF (Franke y Svergun, 2009). Con el fin de mejorar la resolución, se aplicó una simetría P2 tanto para CsdA₂ como para (CsdA-CsdE)₂. Los modelos más representativos fueron comparados con las estructuras cristalográficas conocidas mediante el uso de PyMOL.

Los modelos de CsdA₂ y (CsdA-CsdE)₂ fueron refinados mediante el programa SASREF (Petoukhov y Svergun, 2005), el cual usa un protocolo de alineamiento para generar el ensamblaje de interconexión de las subunidades sin choques estéricos, de tal forma que encaje con los datos de dispersión. En la figura 4.23 se muestran los datos experimentales. En todos los casos los datos cristalográficos eran acordes con los resultados obtenidos mediante SAXS.

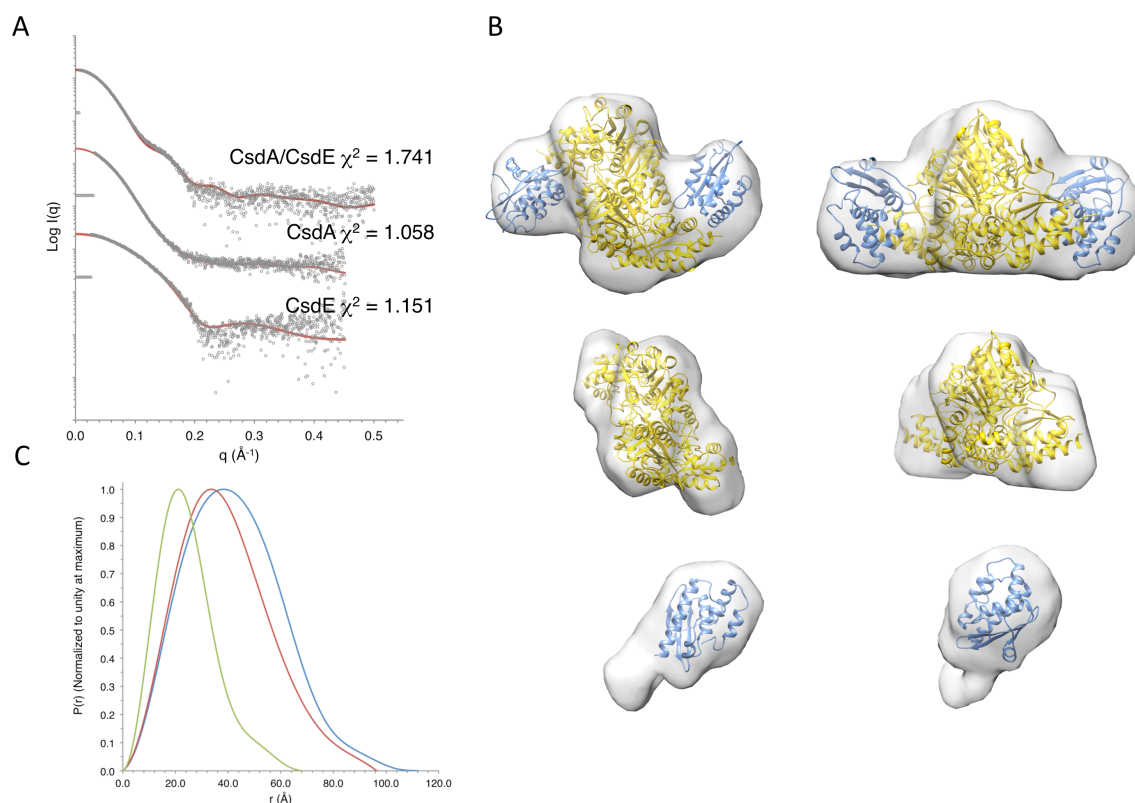


Figura 4.23. Datos de dispersión por SAXS, modelado molecular y formas de (CsdA-CsdE)₂, CsdA₂ y CsdE. **a** Patrones de dispersión de CsdA₂, CsdE y del complejo (CsdA-CsdE)₂. La representación muestra el logaritmo de la intensidad dispersada como una función del momento de transferencia ($q = 4\pi \sin \theta / \lambda$), donde q es el ángulo de dispersión y λ es la longitud de onda del rayo-X. Los patrones experimentales se representan como gráfica de dispersión (puntos grises), mientras que la dispersión teórica para las estructuras de los cristales, se muestran en líneas rojas. **b** Se muestran los mejores modelos para (CsdA-CsdE)₂ (representación superior), así como para los componentes por separado, CsdA₂ (medio) y CsdE (zona inferior), donde la envoltura por SAXS usando DAMMIF (Franke y Svergun, 2009) se muestra en blanco semitransparente, englobando las estructuras cristalográficas en cada uno de los casos. Las estructuras cristalinas están representadas como cintas, mostrando CsdA₂ en dorado y CsdE en azul. **c** Funciones $P(r)$ de distribución de CsdA₂ (rojo), CsdE (verde) y (CsdA-CsdE)₂ (azul). La formación del complejo lleva a un cambio hacia mayores D_{max} para el complejo (CsdA-CsdE)₂.

4.3.6. Cristalización del complejo (CsdAper-CsdEper)₂

Los cristales del complejo persulfurado fueron obtenidos mediante el método de gota sentada por difusión de vapor, aplicando el complejo correspondiente a una concentración de 20 mg/mL (previamente analizado por DLS) a 293 K. La solución de cristalización empleada fue 0,1 M Tris-HCl pH 8,5, 0,1 M acetato de magnesio, 10% (w/v) PEG 8.000. Los cristales aparecieron en 1-4 días. Éstos fueron protegidos con la adición de 30% de glicerol para su congelación en nitrógeno líquido.

4.3.7. Colección y procesamiento de datos de difracción de rayos-X

Los datos fueron colectados en la línea ID23-1 (ESRF, Grenoble, Francia). Los cristales difractados corresponden a un grupo espacial ortorrómbico $P2_12_12_1$. Los cristales difractaron a una resolución de 2,49 Å. En la figura 4.24 se muestra la imagen de difracción de los cristales del complejo (CsdA-CsdE)₂ persulfurado. Los datos obtenidos fueron procesados con XDS (Kabsch, 2010) y escalados gracias a *Aimless* (Evans y Murshudov, 2013).

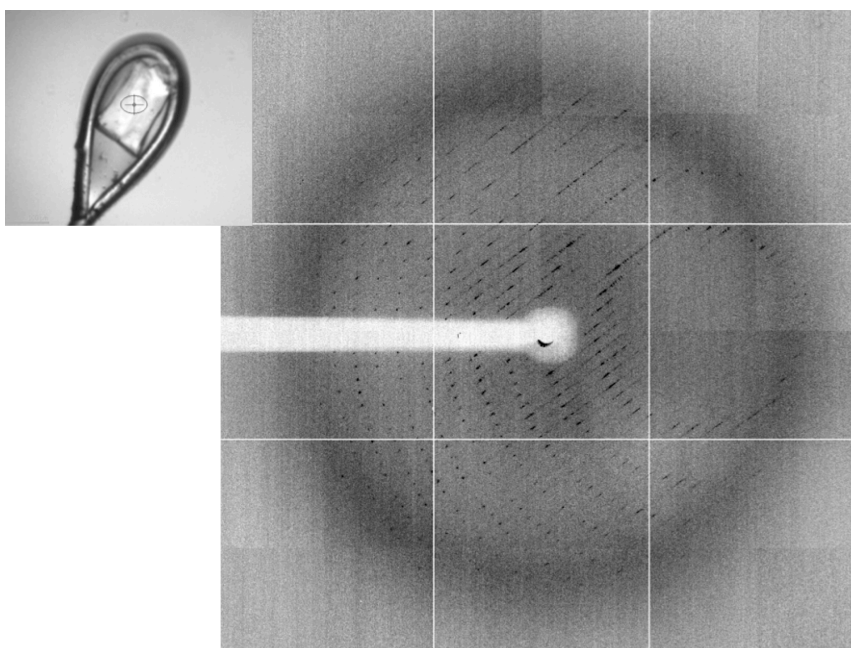


Figura 4.24. Imagen del cristal y de difracción del complejo (CsdA-CsdE)₂ persulfurado. Se muestran a la izquierda la imagen del cristal en el *loop* listo para difractar. A la derecha se muestra la imagen de difracción obtenida.

4.3.8. Resolución estructural y refinamiento del complejo (CsdA-CsdE)₂

La estructura del complejo persulfurado fue resuelta gracias a la aplicación de MR (PHASER (McCoy *et al.*, 2007)). Se emplearon como modelos las estructuras de CsdA₂ (pdb 5ft4) y CsdE (pdb 5ft7) en los procesos de rotación/traslación. La estructura obtenida se refinó manualmente con el uso de *phenix.refine* (Afonine *et al.*, 2012), llevando a cabo una reconstrucción manual y una validación mediante *Coot* (Emsley *et al.*, 2010) y *MolProbity* (Chen *et al.*, 2010). En la tabla 4.5 se muestran las estadísticas.

(CsdA _{per} -CsdE _{per}) ₂	
PDB code	5ft8
Data collection	
Wavelength (Å)	0.98011
Resolution range (Å)	48.3-2.49 (2.63-2.49)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	74.8, 115.1, 604.9
β (°), $\alpha = \gamma = 90^\circ$	90, 90, 90
Total reflections	625,506 (64,000)
Unique reflections	178,722(23,193)
Multiplicity	3.5 (2.8)
Completeness (%)	97.90 (88.30)
Mean I/ σ (I)	7.40 (1.80)
Wilson B-factor	49.53
R-merge	0.127 (0.553)
R-meas ^a	0.1490
CC1/2 ^b	0.989 (0.581)
CC ^{*c}	0.996 (0.815)
Refinement	
R-work	0.2532 (0.3566)
R-free	0.2768 (0.4141)
# non-H atoms	34,325
# Protein	32,886
# Ligand atoms	418
# Water	1021
Protein residues	4296
RMS (bond) (Å)	0.010
RMS(angles)(°)	1.20
Ramachandran analysis	
Favored/Allowed/Outlier (%)	96.0/0.0/0.68
Clashscore	8.09
Average B-factor (Å ²)	68.70
Protein	69.70
Ligands	59.30
Solvent	40.40

Tabla 4.5. Datos estadísticos de la solución de las estructuras del complejo (CsdA_{per}-CsdE_{per})₂.

4.4. Aceptor de azufre del sistema Suf (SufE)

4.4.1. Clonaje de SufE

El gen codificante para la SufE se obtuvo del genoma de *E. coli* (XL1blue) mediante la técnica PCR y el uso de los oligonucleótidos específicos para este gen (anexo I, tabla 1). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes (*NdeI/XhoI*), se clonó en el vector de expresión con resistencia a ampicilina pET21b, previamente digerido con las mismas enzimas. Una vez obtenida la construcción pET21b- SufE, se transformó en la cepa de clonaje de *E. coli* XL1blue, creciendo colonias en placas de LB con agar al 1,5% más ampicilina 100 µg/mL. Una vez comprobado que la construcción era correcta mediante PCR de colonia y secuenciación con los oligonucleótidos T7 promotor y 3' específico de gen (anexo I, tabla 1), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.4.2. Expresión de SufE

Se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Esta prueba se realizó en 1 mL de LB más ampicilina a la concentración mencionada anteriormente. De esta forma las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. Seguidamente el cultivo es inducido mediante la adición de 1 mM IPTG (3 horas 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonicado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 µL del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más ampicilina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin embargo, el tiempo de expresión se llevó a 5 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

4.4.3. Purificación de SufE

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular fue resuspendido en el tampón de lisis, 50 mM NaH₂PO₄ pH 8,0, 300 mM NaCl, 20 mM imidazol, 5 mM β-mercaptoetanol, al que se le añade 10 mM de PMSF. Una vez resuspendido el sedimento celular, éste es sonificado (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. A continuación el lisado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 µm de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de una ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando gradualmente la concentración de imidazol hasta 250 mM, sin cambiar ninguno de los otros componentes del tampón. La proteína eluyó a 125 mM de imidazol (figura 4.24). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 15 kDa, correspondiente a la masa molecular del monómero de SufE (figura 4.25).

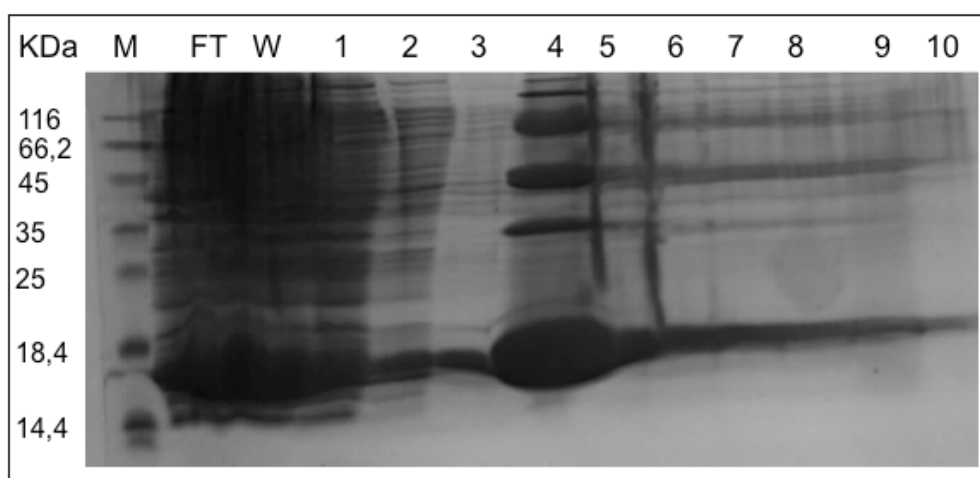


Figura 4.25. Purificación mediante IMAC de SufE. Se muestra el análisis mediante PAGE-SDS de las alícuotas de la purificación, observándose que la mayoría de la proteína, se encuentra sobre 15 kDa, correspondiente a la recombinante SufE. M se refiere al marcador, FT al *flow-through*, W al lavado y los números a las alícuotas analizadas.

Las alícuotas analizadas en el paso anterior mediante PAGE-SDS fueron concentradas mediante el uso de concentradores (*Millipore*), para ser cargadas en una

columna HiLoad 16/60 Superdex200 previamente equilibrada en 10 mM Tris-HCl pH 7,5, 5 mM DTT. La filtración en gel se llevó a cabo a un flujo de 1 mL/min. El proceso de filtración en gel dio un pico de elución de 106,85 mL, correspondiente al monómero de la proteína (figura 4.26). Las muestras de elución fueron analizadas mediante PAGE-SDS.

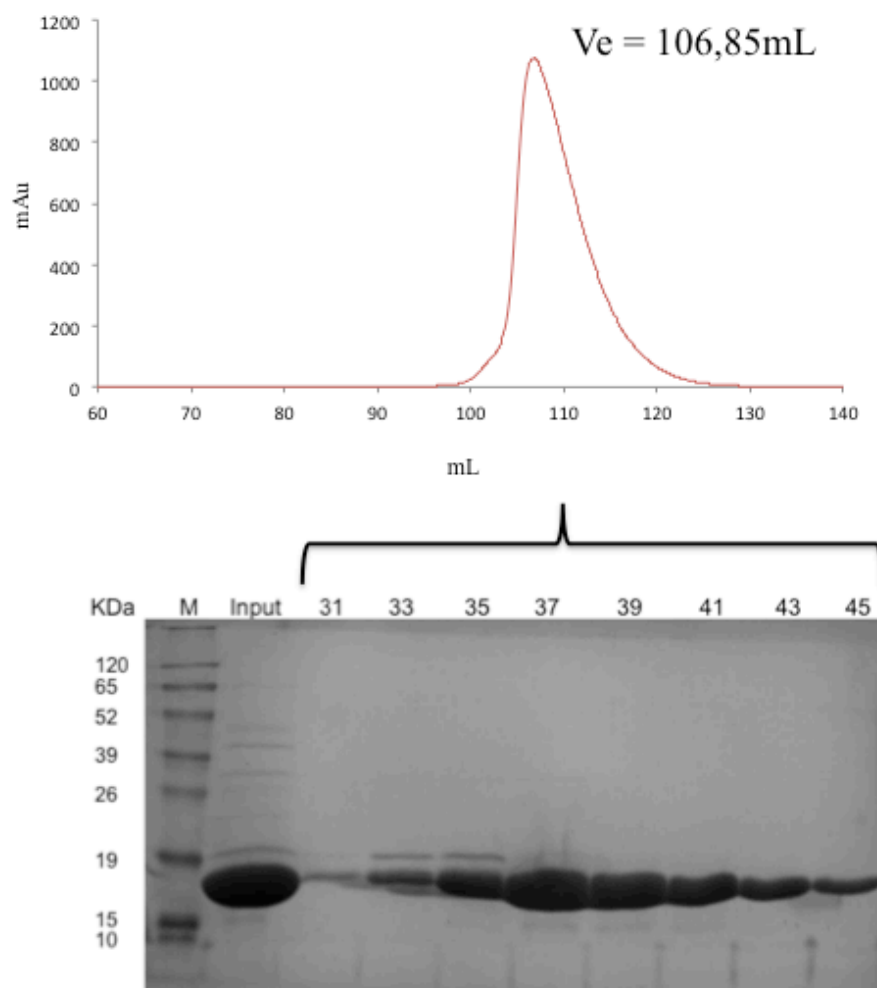


Figura 4.26. Purificación mediante filtración en gel de SufE. Se muestra el cromatograma generado, representando las miliunidades de absorbancia (mAU) a 280 nm frente al volumen. Se puede observar que el volumen de elución (V_e) es de 106,85 mL correspondiente al monómero de la proteína. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M se refiere al marcador, *Input* a la muestra cargada en la columna y los números a las alícuotas analizadas.

4.4.4. Experimentos de transferencia de S entre CsdA, CsdE y SufE (*crosstalk*)

Con el fin de demostrar la existencia de la conexión entre los sistemas de *E. coli* Suf y CSD, se llevaron a cabo experimentos de persulfuración entre algunos de sus componentes, empleando las enzimas CsdA, CsdE y SufE previamente purificadas, y calculando las masas de CsdE y SufE mediante espectrometría de masas tras la reacción de persulfuración.

Todas las proteínas, previamente purificadas, fueron dializadas por separado en el tampón 10 mM HEPES pH 7,5, 150 mM NaCl y 0,1 mM EDTA. De esta forma se incubó CsdA (8 μ M) con 5 mM L-Cys (30 minutos 298 K), la cual fue centrifugada a 10.000 g (10 minutos 277 K), tomando posteriormente el sobrenadante, donde se encuentra CsdA persulfurada. A continuación, ésta fue incubada con 8 μ M de CsdE o SufE (30 minutos 298 K). Seguidamente se llevó a cabo una filtración en gel en la columna Superdex200 10300GL, previamente equilibrada en el tampón mencionado para este experimento, con el fin de aislar el complejo (CsdAper-CsdEper)₂ y las proteínas (SufEper y CsdEper, no capaz de formar el complejo CsdA-CsdE). En cuanto al complejo persulfurado formado por (CsdAper-CsdEper)₂, éste fue incubado con SufE (8 μ M), en el mismo tampón (30 minutos 298 K) sin la adición de L-Cys. Seguidamente se purificaron de nuevo, mediante la filtración en gel mencionada, el complejo y SufE. De esta forma se obtuvieron tanto CsdE como SufE persulfurada mediante dos vías, la primera a través de CsdAper (en los dos casos) y la segunda mediante (CsdAper-CsdEper)₂, solo en el caso de SufE, para a continuación analizarlas mediante espectroscopía de masas (MALDI-TOF) con el fin de evaluar el transporte de azufre. Los resultados de las masas obtenidas en cada caso se muestran en la tabla 4.6. Del mismo modo se llevaron a cabo controles con SufE sin tratar, así como incubada con L-Cys con el fin de descartar que sea capaz de captar por sí misma el azufre del sustrato.

Proteína	Masa molecular (Da)
SufE	16.729,5
SufE + L-Cys	16.731,2
SufEper vía (CsdAper-CsdEper) ₂	16.764,2
SufEper vía (CsdAper) ₂	16.744,2; 16.774,1; 16.809,1
CsdE	16.753,4
CsdEper vía (CsdAper) ₂	16.789,3

Tabla 4.6. Valores de masas del ensayo de espectrometría de masas para la persulfuración de CsdE y SufE. Se muestra las masas moleculares (Da) obtenidas en estos experimentos para los distintos ensayos, así como los controles. En los casos de las persulfuraciones se indica mediante que vía se llevaron a cabo, o (CsdAper-CsdEper)₂ o (CsdAper)₂.

Como se puede observar entre el control negativo de SufE (16.729,5 Da) y su tratamiento con L-Cys (16.731,2 Da) no hay una diferencia significativa, por lo que SufE no es capaz de tomar ese átomo de azufre directamente de la fuente por sí mismo. Por otro lado, cuando se incuban CsdA-CsdE en presencia de L-Cys, y se purifica el complejo, éste sí es capaz de ceder ese átomo a SufE (16.764,2 Da), ya que entre este experimento y el control negativo de SufE (16.729,5 Da) hay una diferencia de 34,7 Da, siendo el valor para el peso molecular de un átomo de azufre de 32 Da.

Por su parte, cuando se incubaba CsdA persulfurada con SufE aparecen tres especies diferentes, 16.744,2, 16.774,1 y 16.809,1 Da, observándose un cambio de 29,9 Da del primero al segundo y de 35 Da del segundo al tercero, lo que hace pensar que nos encontramos ante varias especies de SufE, con 1-3 átomos de azufres incorporados.

4.5. ARNt treonilcarbamoiladenosina deshidratasa (TcdA)

4.5.1. Clonaje de TcdA

El gen codificante para la TcdA se obtuvo del genoma de *E. coli* (XL1blue) mediante la técnica PCR y el uso de los oligonucleótidos específicos para este gen (anexo I, tabla 1). Se emplearon los cebadores para la obtención de una proteína recombinantes con cola de histidinas C-terminal, con el fin de clonar el gen codificante en el vector empleado en este caso (pET21b). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes, se clonó en el vector de expresión con resistencia a ampicilina pET21b. En este caso tanto el vector como el inserto fueron digeridos con las enzimas *NdeI/XhoI*. Una vez obtenida la construcciones pET21b-TcdA, se transformó en la cepa de clonaje de *E. coli* XL1blue, creciendo colonias en placas de LB con agar al 1,5% con ampicilina 100 μ g/mL. Una vez comprobado que la construcción era correcta mediante PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen en cada caso (anexo I, tabla 1), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.5.2. Expresión de TcdA

Posteriormente se seleccionaron dos colonias de BL21(DE3) de cada construcción, con el fin de llevar a cabo las pruebas de expresión, las cuales se realizaron del mismo modo en ambos casos. Estas pruebas se realizaron en 1 mL de LB más ampicilina a la concentración mencionada anteriormente. De esta forma las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación la temperatura se disminuye a 303 K y se induce el cultivo con la adición de 1 mM IPTG (3 horas 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 μ L de PBS pH 7,0, para ser posteriormente sonicado. Tras la lisis celular se tomaron 15 μ L de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 μ L del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más ampicilina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin

embargo, el tiempo de expresión se llevó a 12 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

La expresión de la proteína doblemente marcada (^{13}C , ^{15}N), en el caso de TcdA con cola de histidinas, se realizó del mismo modo cambiando simplemente el medio de expresión. Para ello se empleó el medio mínimo de cultivo M9 (100 mM Na_2HPO_4 , 20 mM KH_2PO_4 , 10 mM NaCl , 25 mM D-glucosa (^{13}C), 20 mM NH_4Cl (^{15}N), 1 mM MgCl_2 , 0,1 mM CaCl_2 y 35 μM tiamina). El resto de los parámetros se respetaron.

4.5.3. Purificación de TcdA con colas de histidinas

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante TcdA con cola de histidinas. Para ello el sedimento celular generado fue resuspendido en el tampón de lisis (20 mM fosfato de sodio/potasio pH 7,4, 300 mM NaCl , 20 mM imidazol y 2 mM β -mercaptoetanol), al que se le añade 10 mM de PMSF. Una vez resuspendido el sedimento celular, éste es sonificado (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. A continuación el lisado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 μm de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de una ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando gradualmente la concentración de imidazol hasta 250 mM, sin cambiar ninguno de los otros componentes del tampón. La proteína eluyó a 125 mM de imidazol (figura 4.27). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 28 kDa, correspondiente a la masa molecular del monómero de TcdA (figura 4.27).

Resultados

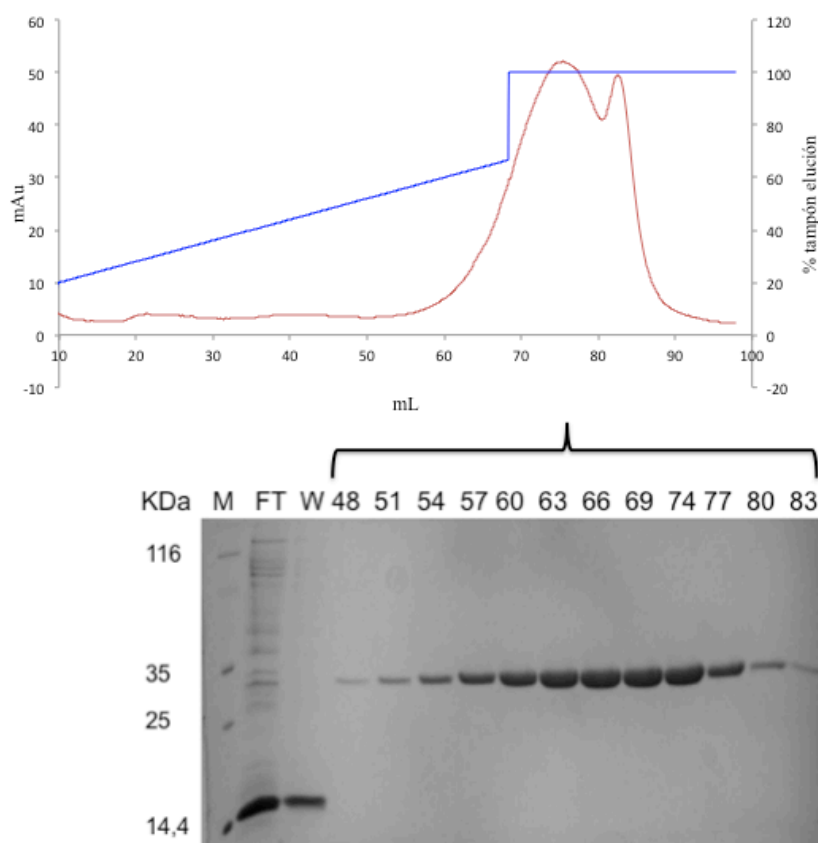


Figura 4.27. Purificación mediante IMAC de TcdA. Se representa el cromatograma de la IMAC mostrándose las miliunidades de absorbancia (mAU) a 280 nm (línea roja) frente al volumen de elución. La línea azul representa la concentración del tampón de elución, observándose el gradiente. En la PAGE-SDS se muestran las muestras de la purificación, observándose que la mayoría de la proteína, sobre 28 kDa, corresponde a la recombinante TcdA. Se puede observar la gran pureza de la muestra. En el gel de PAGE-SDS M es el marcador, FT el *flow-through*, W el lavado y los números corresponden a las alícuotas analizadas.

A continuación la proteína es concentrada mediante el uso de concentradores (Millipore). Tras concentrar la proteína ésta es cargada en una columna de filtración en gel (Superdex75 10/300GL, GE Healthcare), equilibrada previamente con el tampón 20 mM fosfato de sodio/potasio pH 7,4, 300 mM NaCl y 2 mM β -mercaptoetanol. La filtración en gel fue llevada a cabo a 0,5 mL/min, eluyendo la proteína al volumen de elución esperado para el homodímero, 10,70 mL (figura 4.28), de aproximadamente 56 kDa. De nuevo las alícuotas del pico de elución obtenido fueron analizadas mediante PAGE-SDS, donde se puede apreciar el grado de pureza de la proteína.

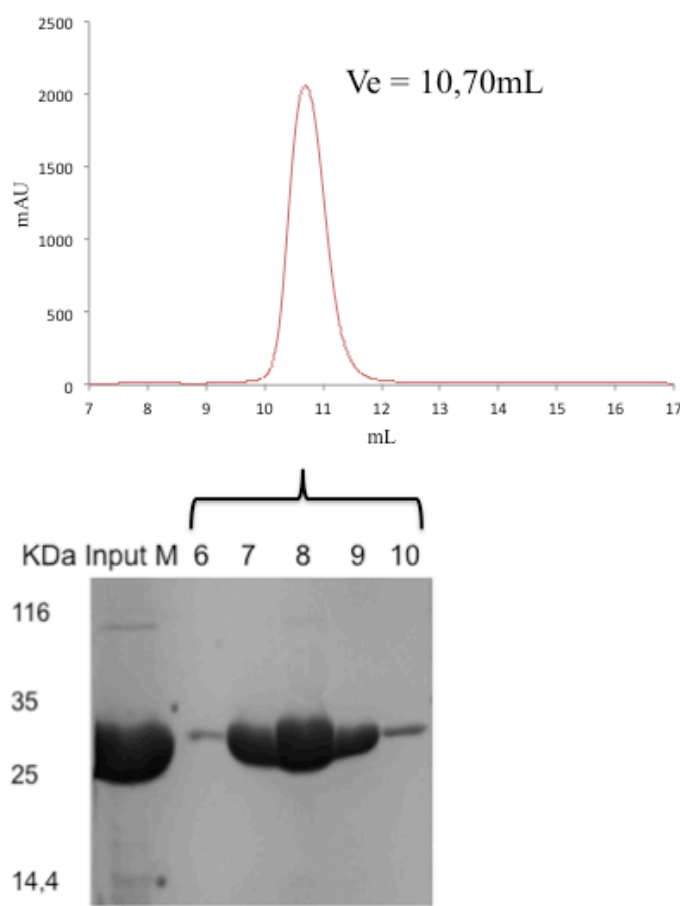


Figura 4.28. Purificación mediante filtración en gel de TcdA. Se representa el cromatograma de la purificación, donde la línea roja muestra la medida de miliunidades de absorbancia (mAU) a 280 nm. Se puede observar que el volumen de elución (V_e) es de 10,70 mL correspondiente al homodímero de la proteína. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En el gel de PAGE-SDS M es el marcador, *Input* es la muestra cargada en la columna y los números corresponden a las alícuotas analizadas.

4.5.4. Clonaje de ARNt^{Lys}(UUU)

El gen codificante para el ARNt^{Lys} (UUU) de *E. coli*, correspondiente al aminoácido lisina fue sintetizado (Genscript) e introducido en un plásmido con resistencia a ampicilina. Este plásmido fue usado como molde para la amplificación del gen del ácido nucleico mediante la técnica de PCR y los oligonucleótidos específicos del gen codificante para el ARNt de la tabla 1 del anexo I. El inserto generado fue digerido con las enzimas *EcoRI*/*XhoI*. Posteriormente se procedió a la ligación en el vector pET23a, digerido con las mismas enzimas. Como en los casos anteriores se transformó en

Resultados

XL1blue, creciendo colonias en placas de LB con agar al 1,5% más ampicilina 100 µg/mL. Tras la obtención de colonias se confirmó la construcción mediante secuenciación, empleando los oligonucleótidos para el T7 promotor y 3' específico de gen (anexo I, tabla 1). Posteriormente se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.5.5. Expresión de ARNt^{Lys}(UUU)

La expresión se realizó en un cultivo de 200 mL de LB más ampicilina (100 µg/mL) en un matraz de 1 L a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. En este momento se induce la expresión con 1 mM de IPTG durante 3 horas a la misma temperatura. El sedimento celular del cultivo se conservó a 195 K.

En esta ocasión además se tomó una alícuota antes y después de la inducción, normalizando la cantidad de células mediante su D.O.₅₉₀, con el fin de comprobar cuál era el fondo de ARNt en el análisis de sobreexpresión, es decir, en qué proporción se encontraba nuestro ARN producido en comparación con el fondo celular. Estas muestras fueron analizadas mediante electroforesis con un gel de agarosa al 2% (figura 4.29).

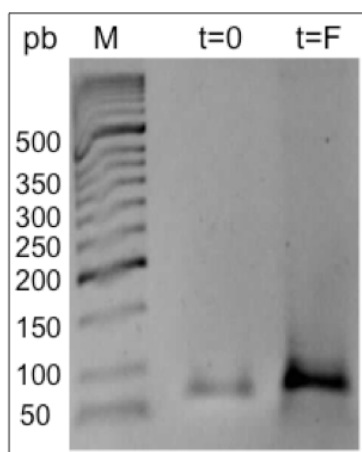


Figura 4.29. Prueba de expresión de ARNt^{Lys}(UUU) analizado mediante gel de agarosa 2%. Se corrieron las muestras antes de inducir (tiempo 0, t=0) y otra representativa del final de la expresión (tiempo final, t=F) con el fin de evaluar el fondo de los ARNt que se encuentran en la célula. pb son los pares de bases y M es el marcador.

De esta forma se observa que la mayoría de las moléculas detectadas de ARNt son correspondientes a la molécula recombinante (75 pb), banda a la altura entre los marcadores de 50 y 100 pb.

4.5.6. Purificación de ARNt^{Lys}(UUU)

Todos los tampones empleados en esta purificación fueron previamente autoclavados, con el fin de evitar la acción de ARNasas presentes en las soluciones. Además siempre que se pudo se mantuvieron las muestras con ARNt a 277 K. El primer paso de la purificación consistió en resuspender el sedimento celular en 5 mL del tampón de lisis (0,3 M acetato de sodio pH 4,5, 10 mM EDTA), añadiendo 1 volumen de fenol ácido (pH 4,3) en último lugar. La mezcla se agita durante un minuto para a continuación centrifugarla a 10.000 g (10 minutos 277 K), para posteriormente tomar la fase acuosa, a la cual se le añade 2,5 volúmenes de etanol al 100%. La mezcla es agitada e incubada (1 hora 277 K) para después ser centrifugada a 20.000 g (20 minutos 277 K). El sedimento resultante es resuspendido en 20 mM sodio/potasio de fosfato pH 7,4, 0,1 mM EDTA, para ser cargado en una columna HiLoad 16/60 Superdex75 (*GE Healthcare*) conectada al equipo de HPLC, previamente equilibrada con el mismo tampón en el que se encuentra la muestra. La filtración en gel se llevó a cabo a 1 mL/min. El ARNt eluyó con un volumen de elución de 75,53 mL, generándose un cromatograma (figura 4.30), lo que permitió analizar las diferentes alícuotas en un gel de agarosa al 2%. Las muestras seleccionadas fueron guardadas a 253 K para su posterior tratamiento.

Resultados

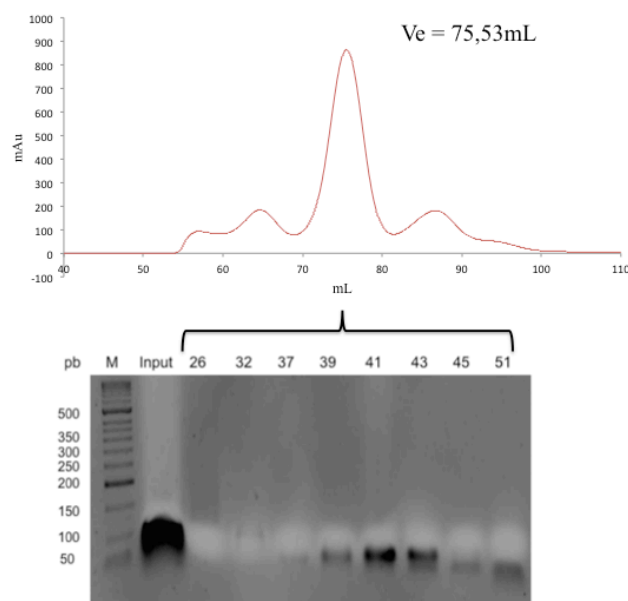


Figura 4.30. Purificación mediante filtración en gel de ARNt^{Lys}(UUU). Se muestra el cromatograma donde se representa las miliunidades de absorbancia (mAU) a 254 nm frente al volumen de elución, obteniéndose un volumen de elución (V_e) de 75,53 mL para el pico correspondiente al ARNt recombinante. El gel de agarosa 2% muestra el análisis correspondiente al pico del ARNt expresado, teniendo el tamaño esperado. pb corresponde a los pares de bases, M al marcador e *Input* a la muestra cargada en la columna y los números a las alícuotas analizadas.

4.6. Estudio cristalográfico de TcdA en complejo con ATP y AMP

4.6.1. Cristalización de TcdA en complejo con ATP y AMP

Los cristales de TcdA acomplejada con ATP o AMP se consiguieron mediante el método de gota sentada (difusión de vapor). Para ello se empleó la proteína purificada a 5 mg/mL (previamente analizada por DLS), incubándola (30 minutos 293 K) con 10 mM de ATP o con 10 mM de AMP más 10 mM de $MgCl_2$ en ambos casos (cocrystalización). La solución de cristalización empleada fue 0,05 M de fosfato de potasio y 10% (v/v) PEG 1.000 a pH básico. Tras 5 días se formaron los cristales. Se añadió glicerol al 20% con el fin de congelarlos en nitrógeno líquido.

4.6.2. Colección y procesamiento de datos de difracción de rayos-X de los cristales de TcdA en complejo con ATP y AMP

Se llevaron a cabo dos colecciones de datos de los cristales obtenidos del complejo TcdA-ATP, el primero de ellos a la máxima resolución ($1,77 \text{ \AA}$) a una longitud de onda de $0,97948 \text{ \AA}$. La segunda colección alcanzó una resolución de $2,35 \text{ \AA}$ a una longitud de onda de $1,99976 \text{ \AA}$ para la difracción anómala del azufre (SAD, *Sulfur Anomalous Diffraction*) en la línea BL13-XALOC (ALBA, Barcelona, España). En la figura 4.31 se muestra una de las imágenes de difracción de los cristales de TcdA acoplada con ATP.

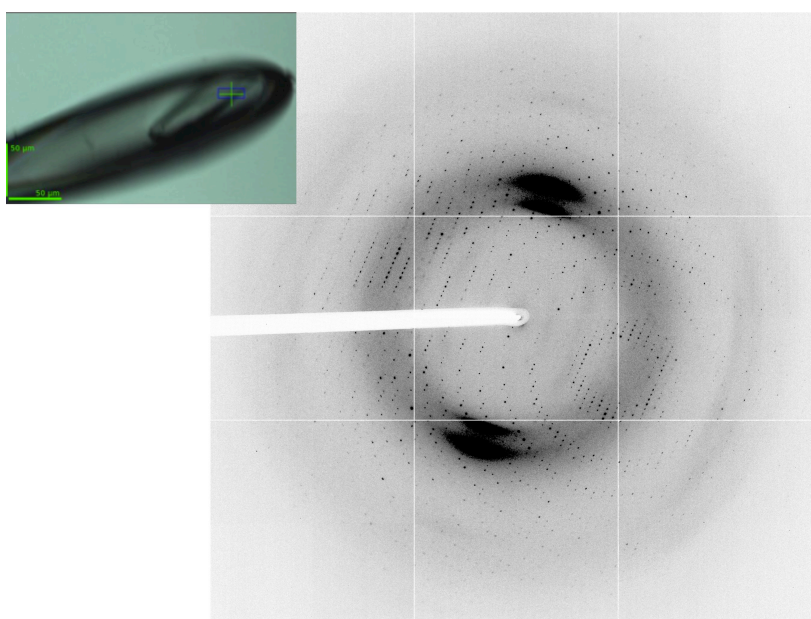


Figura 4.31. Imagen de difracción de un cristal de TcdA acoplado con ATP. Se muestran a la izquierda la imagen del cristal en el *loop* listo para difractar. A la derecha se muestra la imagen de difracción obtenida.

En cuanto al complejo formado con AMP, se colectaron datos de éste a la resolución $1,80 \text{ \AA}$ en la línea PROXIMA 2 (SOLEIL, París, Francia). Todos los datos obtenidos fueron integrados con XDS (Kabsch, 2010) y escalados con la ayuda de *Aimless* (Evans y Murshudov, 2013).

4.6.3. Resolución estructural y refinamiento de TcdA en complejo con ATP y AMP

La estructura de TcdA-ATP fue resuelta a una resolución de 1,77 Å, mediante el uso de MR-SAD, empleando como modelo la proteína MoeB de *E. coli* (PDB 1jwa), en concreto los residuos 21-199, mediante el uso de PHASER (McCoy *et al.*, 2007). A continuación, gracias al modelo generado, se usó la señal anómala del azufre, del fósforo, proveniente de la molécula de ATP y del potasio. De esta forma la estructura completa de la proteína TcdA pudo ser construida en base al mapa de densidad electrónica generada desde las fases del MR-SAD. Sin embargo, la zona entre los residuos 217-236, por encima del sitio de unión al ATP, no pudo ser resuelta debido a su flexibilidad.

La estructura de la proteína acomplexada con AMP fue resuelta mediante MR, empleando como modelo la estructura TcdA-ATP. Ambas estructuras fueron construidas, refinadas y validadas con Coot (Emsley *et al.*, 2010), phenix.refine (Afonine *et al.*, 2012) y MolProbity (Chen *et al.*, 2010). En la tabla 4.7 se muestran las estadísticas obtenidas de estos procesos.

	TcdA-ATP	TcdA-AMP
PDB code	4d79	4d7a
Data collection		
Wavelength (Å)	0.9795	0.9801
Resolution range (Å)	41.13-1.77 (1.83-1.77)	41.14-1.80 (1.86-1.80)
Space group	$P 2_1$	$P 2_1$
a, b, c (Å)	65.3, 96.7, 82.8	65.7, 97.2, 83.2
β (°), $\alpha = \gamma = 90^\circ$	90, 111.2, 90	90, 111.6, 90
Total reflections	312,887 (11,351)	351,361 (33,129)
Unique reflections	89,883 (5907)	87,219 (8261)
Multiplicity	3.5 (1.9)	4.0 (4.0)
Completeness (%)	95.96 (63.90)	97.14 (92.68)
Mean I/σ (I)	18.97 (3.75)	19.93 (4.05)
Wilson B-factor	23.95	21.59
R-merge	0.08388 (0.663)	0.04687 (0.412)
R-meas^a	0.09797	0.05407
CC1/2^b	0.995 (0.556)	0.999 (0.875)
CC^c	0.999 (0.846)	1.000 (0.966)
Refinement		
R-work	0.1416 (0.2994)	0.1396 (0.1839)
R-free	0.1833 (0.3236)	0.1768 (0.2350)
# non-H atoms	8167	8064
# Protein	7484	7411
# Ligand atoms	172	46
# Water	511	607
Protein residues	996	974
RMS (bond) (Å)	0.011	0.007
RMS(angles)(°)	1.430	1.01
Ramachandran analysis		
Favored/Allowed/Outlier (%)	98.0/2.0/0.0	98.0/2.0/0.0
Clashscore	1.88	1.93
Average B-factor (Å²)	34.40	32.10
Protein	33.60	31.50
Ligands	60.80	49.70
Solvent	38.00	37.20

Tabla 4.7. Datos estadísticos de la solución de las estructuras de TcdA acomplejada con ATP y AMP.

Se pudo construir un modelo tridimensional del dímero de TcdA, confirmado su pertenencia, en cuanto al plegamiento, a la superfamilia de las enzimas de tipo E1. La estructura tridimensional del monómero de TcdA está compuesta de 7 hebras β , localizadas en una lámina rodeada de 8 hélices α . Todas las hebras β son paralelas a excepción de β_6 , situada al final de la lámina. Como se puede observar en otros componentes de la superfamilia MoeB/E1, la mitad N-terminal de la lámina descrita contiene una variación del plegamiento Rossmann, típico de proteínas capaces de unir nucleótidos, encontrándose la topología $\beta\alpha\beta\alpha\beta$ interrumpida entre β_2 y α_4 por una inserción de dos hélices de tipo 3_{10} consecutivas. La primera de estas hélices ($H3_{10a}$) contiene tres residuos conservados en la superfamilia MoeB/E1, incluyendo la Asn69, encargada de interaccionar con la molécula de ATP o con la de AMP, mediante puentes de hidrógeno mediados por moléculas de agua. En cuanto a la capacidad de hidrólisis del ATP, el *loop* entre β_1 y α_3 contiene el motivo altamente conservado G-X-G-A/G-L/I-G, donde X representa cualquier aminoácido, típicamente valina, leucina o isoleucina (figura 4.32).

La mitad C-terminal de TcdA está compuesta por 3 hebras β (β_4 , β_5 y β_7), unidas mediante puentes de hidrógeno a la hebra antiparalela β_6 , dos hélices α conectando consecutivamente con hebras β (α_6 y α_7) y una larga α hélice (α_8), que se prolonga perpendicularmente en dirección al eje de simetría del dímero. Varios motivos en la mitad C-terminal, incluyendo β_6 , α_7 y β_7 , así como los *loops* que conectan entre ellos generan dos sitios de unión a metales, uno para la unión al K^+ (por cada monómero de la proteína) y otro para la unión del Na^+ (en la zona de dimerización) (figura 4.32).

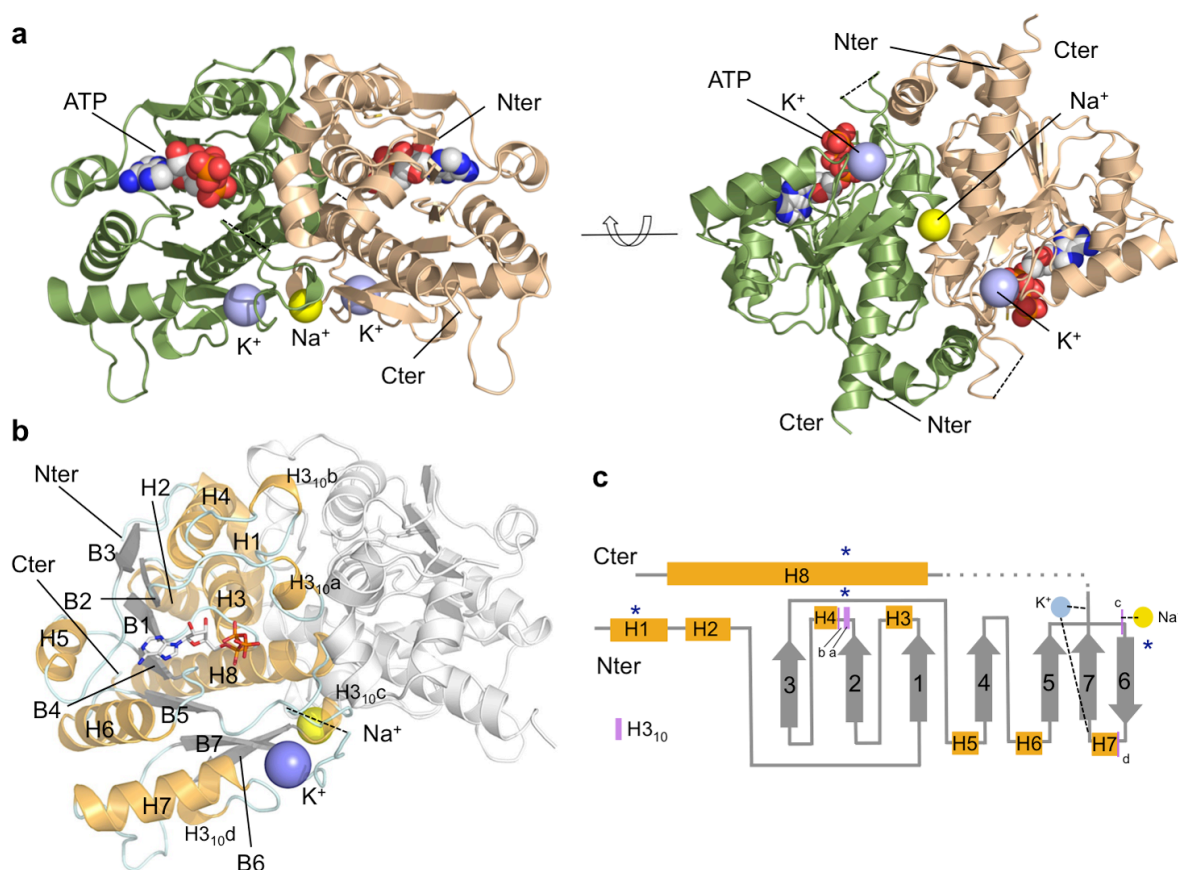


Figura 4.32. Estructura resuelta por cristalografía de rayos-X de TcdA a 1,77 Å de resolución en complejo con ATP. **a** Representación de cintas del dímero de TcdA en dos orientaciones rotadas 90° una respecto de la otra, representando cada subunidad en un color diferente (verde y color trigo). Dos átomos de K⁺ y uno de Na⁺ se muestran como esferas violetas y amarilla, respectivamente. El catión de Na⁺ se localiza en la zona de dimerización y su coordinación compromete a residuos de ambas subunidades, así como a moléculas de agua. La molécula de ATP está unida a una zona conservada, mostrada en colores CPK. **b** Representación de la estructura secundaria de los diferentes elementos del monómero de TcdA. Las hélices se muestran en dorado (hélices α marcadas como H1-H8 y la de tipo 3₁₀ como H3₁₀ a-d), las hebras β en blanco (B1-B7), y los *loops* en cian claro. **c** Anotación esquemática de la topología de TcdA. Las hélices y las hebras están marcadas como en (b), excepto en el caso de las hélices de tipo 3₁₀, indicadas con líneas púrpuras y marcadas de a-d. Los iones de K⁺ y Na⁺ se muestran como círculos azules y amarillos, respectivamente. Las líneas discontinuas delimitan las regiones que interaccionan con K⁺, y las estructuras involucradas en la dimerización de TcdA se marcan con asteriscos azul oscuro.

En la estructura de TcdA resuelta con la molécula de ATP, se observa que esta molécula se encuentra en un hueco creado entre el extremo C-terminal y la lámina β

Resultados

central. Los residuos del *loop* P forman la base del motivo trifosfato del nucleótido, localizando el anillo de adenina en una cavidad hidrofóbica. Interacciones específicas anclan la molécula de ATP al sitio de activación a través del motivo trifosfato de esta molécula, además de los grupos hidroxilo de la ribosa. Un puente de hidrógeno entre la Gly40 (en el *loop* P) mantiene el grupo α -fosfato del ATP unido al bolsillo en una conformación torcida. Los residuos conservados Arg72 y Lys85 interaccionan con los oxígenos de los fosfatos α y β , respectivamente, mientras que la Thr68 y la Asn69 en la hélice H3₁₀ anclan el fosfato γ mediante un puente de hidrógeno.

En cuanto al complejo TcdA-AMP, el nucleótido se estabiliza de forma análoga al explicado en el caso anterior, sin observarse en la estructura cambios conformacionales. En la figura 4.33 se muestran las zonas de interacción de TcdA con ATP y AMP.

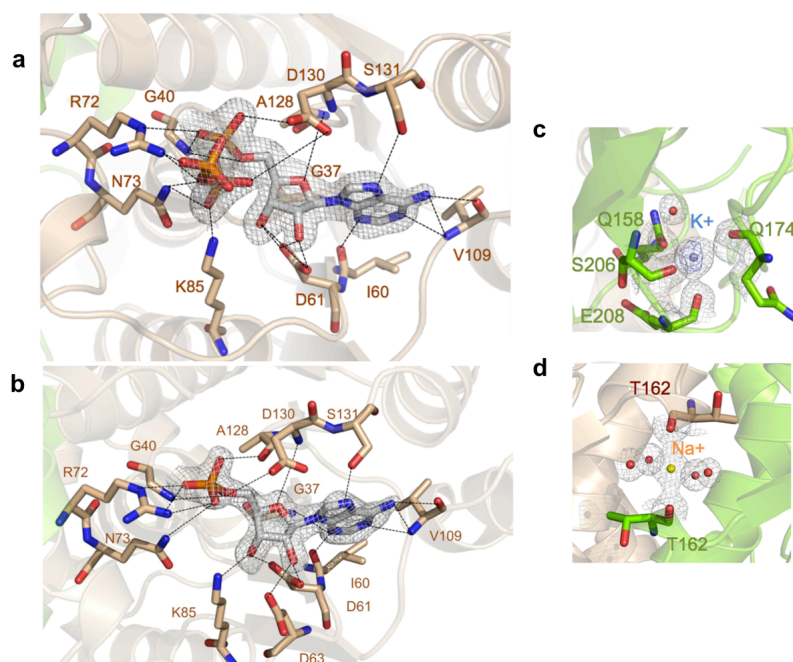


Figura 4.33. Mapas de densidad de electrones mediante rayos-X de los bolsillos de unión a moléculas de ATP y AMP en el centro activo de TcdA, así como las esferas de coordinación de los iones de K⁺ y Na⁺. **a** Vista detallada del sitio de unión de ATP en TcdA. Los residuos que interaccionan con esta molécula se marcan y se muestran como barras. El mapa de densidad electrónica se representa alrededor de la molécula de ATP a un nivel de contorno de $1,5 \sigma$. **b** Vista detallada de la molécula de AMP en el centro activo de TcdA. Representación y mapa de densidad electrónica con las mismas características que en el caso de **(a)**. **c** y **d** Esferas de coordinación de K⁺ **(c)** y Na⁺ **(d)**. Los residuos de interacción están marcados y se muestran como barras. Las moléculas de agua se marcan como esferas rojas. Los metales y las esferas de coordinación se muestran con la densidad electrónica (gris). El mapa anómalo de Fourier calculado de los datos obtenidos a longitud de onda (1,89 Å) a una resolución de 2,4 Å se muestra para el caso de K⁺ **(c)** en púrpura. Para el caso del Na⁺ no hay señal anómala a esa longitud de onda, facilitando la discriminación entre K⁺ y Na⁺ en la estructura.

En ambos complejos (tanto con ATP como con AMP), la región entre $\beta 7$ y el C-terminal de $\alpha 8$, la cual abarca los residuos 214-236, se encuentra totalmente desordenada en los mapas de densidad electrónica, por lo que no es posible trazarlo en la estructura. Este *loop* desordenado, muy próximo al sitio de unión al ATP, corresponde con la zona de desorden de la proteína MoeB (Lake *et al.*, 2001), por lo que es fácil pensar que pueda tener alguna función catalítica o alguna función en la comunicación proteína-proteína.

Resultados

Siguiendo con la estructura de TcdA, uno de los bolsillos de unión al metal K^+ en cada uno de los monómeros se encuentra enterrado, tal y como muestran los mapas anómalos de Fourier de este ion. La geometría de la esfera de coordinación de K^+ es de tipo tetrahédrica distorsionada. Los residuos involucrados incluyen los oxígenos del carbonilo de la cadena principal en el caso de Gln174, Ser206 y Glu208, así como la cadena lateral de Gln158. El sitio de unión del K^+ está creado por los elementos estructurales de la mitad C-terminal del monómero, comprendiendo los *loops* $\beta 6$ - $\alpha 7$, $\beta 7$ y el *loop* $\beta 7$, siendo la única superficie ausente en otros miembros de la superfamilia MoeB/E1, por lo que podría albergar funciones específicas de esta proteína, en cuanto a unión/modificación de ARNt u otras.

El segundo sitio de unión para el otro metal, el Na^+ , se encuentra en la zona de dimerización en una esfera de coordinación octaédrica. Este catión está axialmente ligado por los oxígenos de dos residuos simétricos de Thr162, que reposan en el segmento de la H3_{10C} del monómero de la otra cadena, con cuatro moléculas de agua ocupando las posiciones ecuatoriales. En la figura 4.33 se muestra la zona de interacción de los dos metales.

Por otro lado, la estructura resuelta por cristalografía de rayos-X muestra una simetría dimérica, consistente con los resultados obtenidos mediante filtración en gel y los datos de SAXS.

La interfaz de dimerización tiene aproximadamente una superficie de 2.250 Å², implicando más del 25% de los residuos de la proteína. De acuerdo con PISA (Krissinel E, 2011), el dímero de TcdA es termodinámicamente estable como indica su energía libre de Gibbs de disociación (-21,5 kcal/mol). Esta superficie de dimerización se compone de cuatro hélices α ($\alpha 1$, $\alpha 3$, $\alpha 4$ y la hélice C-terminal $\alpha 8$, de cada monómero), las hélices H3_{10a} y H3_{10C}, que forman el sitio de unión a nucleótidos y al Na^+ respectivamente, y los *loops* $\beta 5$ - $\beta 6$ y $\beta 7$ - $\alpha 8$. Las interacciones que estabilizan la interfaz dimérica incluyen enlaces de hidrógeno, puentes salinos e interacciones hidrofóbicas. Así se ven involucrados residuos de arginina superficiales localizados en las hélices $\alpha 1$ (Arg12), $\alpha 3$ (Arg51), H_{10b} (Arg78) y $\alpha 4$ (Arg42). En la figura 4.34 se muestran los residuos y estructuras de TcdA involucrados en la dimerización.

4.7. Estudio de las interacciones de TcdA con el sistema CSD (CsdA y CsdE)

4.7.1. Caracterización de la unión de TcdA y CsdA por RMN

Con el fin de estudiar las posibles interacciones entre TcdA y CsdA se llevaron a cabo ensayos mediante la técnica RMN, empleando TcdA previamente marcada con

Resultados

los isótopos ^{13}C y ^{15}N . Para ello se llevaron a cabo espectros de tipo 1D ^1H - ^{13}C y ^1H - ^{15}N HSQC, empleando 50 μM de TcdA doblemente marcada y un exceso molar 10 veces superior de CsdA previamente purificada. El resultado fue comparado con una muestra donde solo se analizó ^{13}C , ^{15}N TcdA como control. De esta forma se observó una disminución de la intensidad de la señal de la resonancia de TcdA en presencia de CsdA, sugiriendo la existencia de una interacción débil entre estas dos proteínas, aunque detectable. Sin embargo, con el ratio 1:10 se sigue observando mucha TcdA libre, por lo que no es una interacción fuerte (figura 4.35).

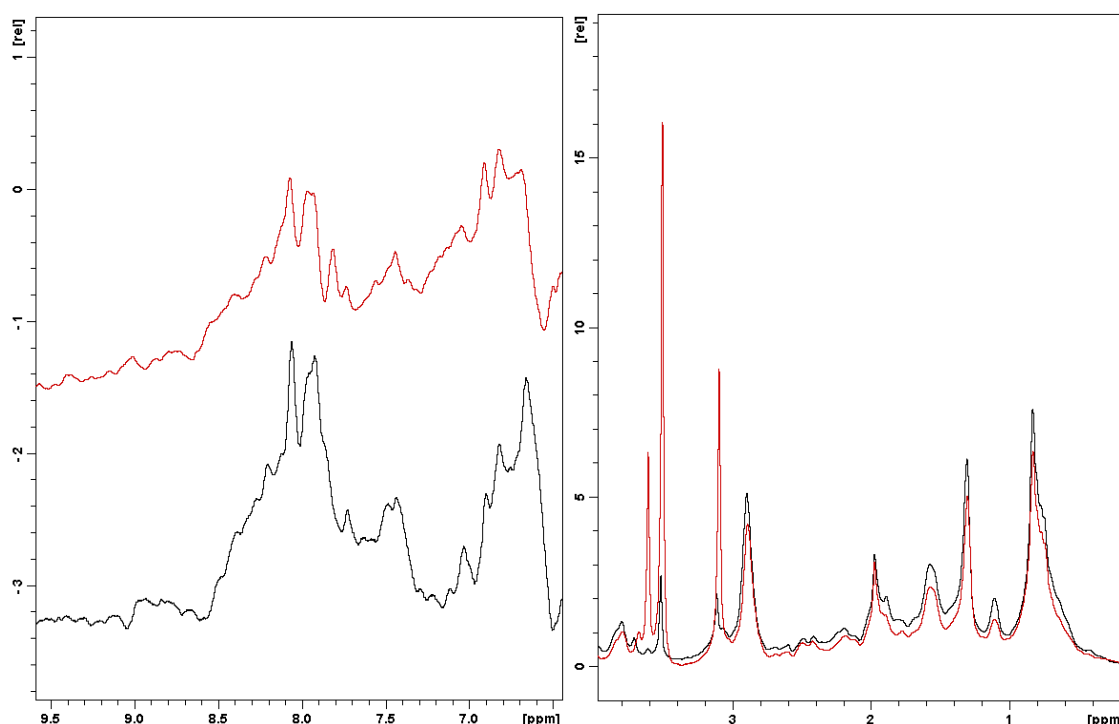


Figura 4.35. Espectros 1D ^1H - ^{13}C y ^1H - ^{15}N HSQC de TcdA doblemente marcada en su interacción con CsdA. En negro ^{13}C , ^{15}N -TcdA, en rojo, ^{13}C , ^{15}N -TcdA en combinación con CsdA. En el panel de la izquierda: espectro de 1D ^1H - ^{15}N HSQC en la región amida/aromática. En el panel de la derecha: espectro de 1D ^1H - ^{13}C HSQC en la región alifática. En el eje vertical se representa la intensidad de la señal mientras que el eje horizontal se representan el cambio químico en ppm (partes por millón). Se observa que en el caso del experimento 1D ^1H - ^{15}N HSQC es donde realmente se observan diferencias, no siendo tan destacables en el caso del experimento de 1D ^1H - ^{13}C HSQC.

4.7.2. Caracterización de la unión de TcdA y CsdE por RMN

Del mismo modo que en el apartado anterior, se llevaron a cabo ensayos de la interacción entre TcdA y CsdE, marcando en este caso solamente la última de ellas (^{13}C , ^{15}N CsdE). Se facilitó mucho el ensayo en este caso, por ser CsdE una proteína relativamente pequeña (15 kDa) y por estar resuelta su estructura mediante la técnica de RMN (bmr5630, PDB 1ni7). Se llevaron a cabo experimentos de tipo ^1H - ^{15}N 2D-HSQC (NOESY). Se generó un espectro con 90 μM de CsdE doblemente marcada y un exceso molar cuatro veces mayor de TcdA sin marcar. Se detectaron cambios químicos en CsdE marcada, tan solo en presencia de TcdA, midiendo por lo tanto una interacción transitoria entre estas dos proteínas.

A continuación, gracias a la información de la interacción entre TcdA y CsdE se pudo mapear la superficie de contacto en CsdE. Se observó que la zona que sufre mayor perturbación se encuentra entre los residuos 70-81 de CsdE, perteneciendo a la horquilla previa a la cisteína catalítica de CsdE (Cys61). En la figura 4.36 se muestran los resultados.

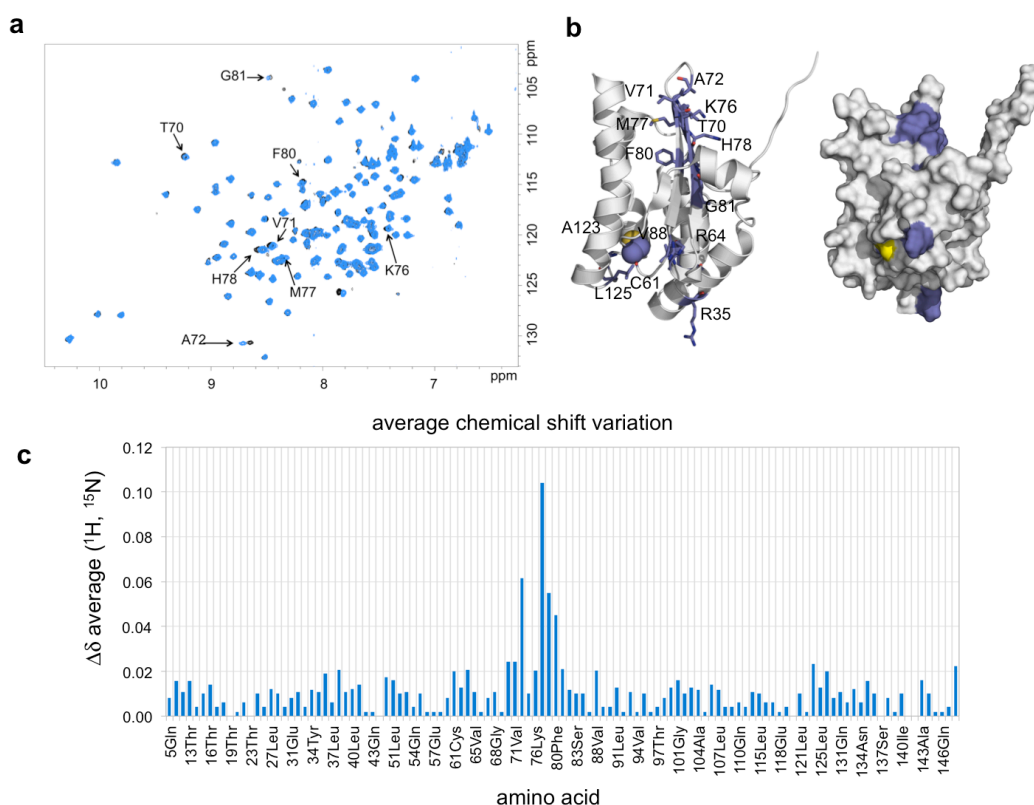


Figura 4.36. Estudios mediante RMN de la interacción transitoria entre TcdA y CsdE. a Espectro de RMN ^1H - ^{15}N NOESY con la ^{13}C , ^{15}N CsdE doblemente marcada en presencia (azul) o ausencia (negro) de TcdA sin marcar. Los cambios significativos en las posiciones específicas de los picos de resonancia de CsdE se encuentran marcados. ppm se refiere a partes por millón de las señales obtenidas de ^1H y de ^{15}N . **b** La posición de los residuos de los aminoácidos afectados se mapean sobre la estructura de RMN de CsdE (pdb 1ni7) en representación de tipo cintas (izquierda) o de tipo superficial (derecha). **c** Media del cambio químico ($\Delta\delta$) de ^{13}C , ^{15}N CsdE sobre la unión a TcdA como medida del espectro NOESY en una representación contra los residuos de la secuencia de CsdE, observándose la zona con mayor cambio en CsdE (residuos 70-81).

Como se ha mencionado los cambios más destacados se observan entre los residuos 70-81, mostrados en la figura 4.36 (b y c), zona cercana a la Cys61. Debido a este cambio detectado es posible que se produzca una mayor exposición del residuo catalítico con el fin de interactuar con TcdA.

4.7.3. Caracterización de la unión de TcdA y CsdE por entrecruzamiento

El uso del entrecruzador BMOE fue esencial para la captación de la interacción transitoria entre TcdA y CsdE. Para ello se mezclaron en un ratio molar 1:1 80 μ M de ambas proteínas en el tampón 20 mM fosfato Na/K pH 7,4, 300 mM NaCl, 0,5 mM EDTA. A continuación se realizó una incubación con 1 mM de TCEP (30 minutos 298 K) con el fin de reducir todos los grupos sulfhidrilos de las proteínas, suplementando posteriormente la muestra con 0,2 mM de BMOE (1 hora 298 K), para a continuación analizar la muestra mediante PAGE-SDS (del mismo modo en las purificaciones mostradas). Igualmente se añadieron los controles en los que se incubo BMOE o solo con TcdA o solo con CsdE. En la figura 4.37 se muestra el resultado del análisis mediante PAGE-SDS, donde se puede observar el entrecruzamiento entre estas dos proteínas, TcdA y CsdE necesariamente mediante sus grupos sulfhidrilos.

El mismo procedimiento fue llevado a cabo entre los pares CsdA/CsdE, siendo otro de los controles del experimento debido al previo conocimiento, CsdA/TcdA, dando un resultado negativo en cuanto a la unión, y el par TcdA/ARNt^{Lys}(UUU), dando de nuevo un resultado negativo en cuanto a la unión.

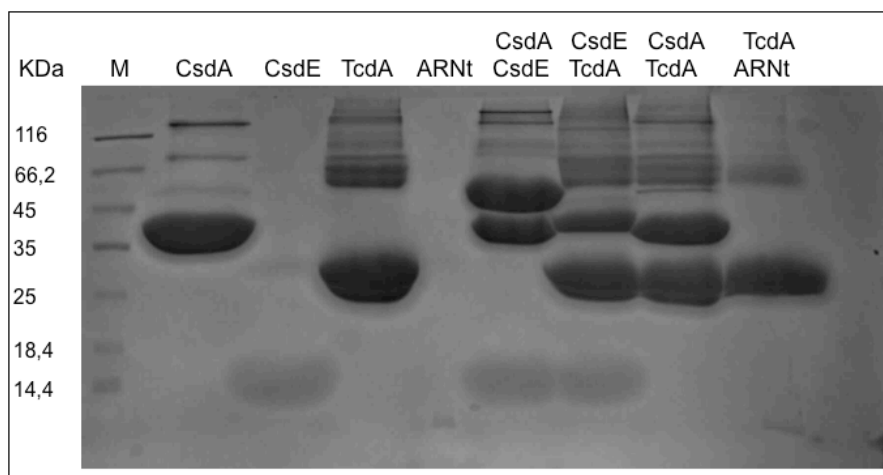


Figura 4.37. Análisis mediante PAGE-SDS del entrecruzamiento mediante BMOE. Los 4 primeros carriles son los controles negativos, tratados con BMOE también. En el resto de los carriles se muestra los pares ensayados, observándose solo en el caso de TcdA-CsdE la aparición de una nueva especie correspondiente al complejo entrecruzado entre un monómero de TcdA y otro de CsdE, a la altura de 43 kDa aproximadamente.

Resultados

Se observa como solo en el caso de TcdA/CsdE aparece una nueva banda a la altura de 43 kDa aproximadamente, siendo el complejo formado entre un monómero de TcdA (28 kDa) y un monómero de CsdE (15 kDa). En el resto de los pares no se ve ningún tipo de entrecruzamiento, salvo el control CsdA/CsdE. Posteriormente este tratamiento se repitió con las proteínas CsdE y TcdA, así como ambas sin tratar como controles negativos para analizarlos posteriormente mediante una PAGE-SDS y MALDI-TOF. Las bandas correspondientes a los monómeros de CsdE y TcdA, que se emplearon como controles, y la banda correspondiente al complejo CsdE-TcdA (compuesto en este caso por un monómero de cada una de ellas) fueron extraídas del gel de acrilamida, para posteriormente ser digeridas con tripsina bajo condiciones no reductoras y analizadas mediante LC-MS (tabla 4.8).

CsdE (secuencia)	MW calculado (Da)
AQAKEIAGC ⁶¹ ENRVWLGYTVAENGK	2.606,3
AQAKEIAGC ⁶¹ ENR	1.289,63
EIAGC ⁶¹ ENRVWLGYTVAENGK	2.209,07
EIAGC ⁶¹ ENR	890,4
AQAKEIAGC ⁶¹ ENR	1.346,65
TcdA (secuencia)	MW calculado (Da)
GKLGVDC ²⁰³ VFSTEALVYPQSDGTVC ²²⁰ AMK	2.890,37
LGVDC ²⁰³ VFSTEALVYPQSDGTVC ²²⁰ AMK	2.705,25
GKLGVDC ²⁰³ VFSTEALVYPQSDGTVC ²²⁰ AMK	2.890,37
LGVDC ²⁰³ VFSTEALVYPQSDGTVC ²²⁰ AMK	2.705,25

Tabla 4.8. Fragmentos generados del complejo CsdE-TcdA (BMOE) mediante digestión por tripsinización. Se muestran los fragmentos que encajarían con los pesos obtenidos, tras la unión covalente del BMOE a las cisteínas que interaccionan entre las dos proteínas y la digestión por tripsinización. En la segunda columna se muestra su masa molecular (MW) en daltons (Da). En el caso de CsdE aparece la única cisteína posible (Cys61), mientras que en el caso de TcdA se identifican dos posibles cisteínas (Cys203 y Cys220).

Se confirmó que el residuo de CsdE Cys61, ya que es la única cisteína que contiene, interacciona o con la Cys203 o con la Cys220 de TcdA, a través de la molécula de BMOE, descartando el resto de las cisteínas encontradas en la secuencia.

4.7.4. Caracterización de la unión de TcdA y CsdE por SAXS

Con el fin de caracterizar el complejo formado por TcdA y CsdE, mediante el entrecruzador BMOE, se llevaron a cabo ensayos del complejo previamente formado, tal y como se detalla en el apartado anterior. Para ello se formó hasta 80 μ M del complejo. Una vez analizado mediante PAGE-SDS, éste fue purificado mediante una filtración en gel en la columna Superdex200 10/300GL, previamente equilibrada en el tampón 20 mM fosfato Na/K pH 7,4, 300 mM NaCl, 0,5 mM EDTA. El pico obtenido de elución fue analizado mediante SAXS (en la línea BM29 BioSAXS (ERSF, Grenoble, Francia)). Los análisis en ScÅtter revelaron un grado de flexibilidad significativo en el complejo, comportándose como una partícula compacta y bien plegada según el exponente de Porod ($P_x = 2,3$). Se aplicó una simetría P2 al modelo, observándose que las subunidades unidas al dímero de TcdA pueden ocupar una posiciones pseudosimétricas a través del dímero de TcdA. En la figura 4.38 se muestra el modelo generado del complejo en base a los resultados de SAXS y las estructuras cristalográficas de ambas proteínas.

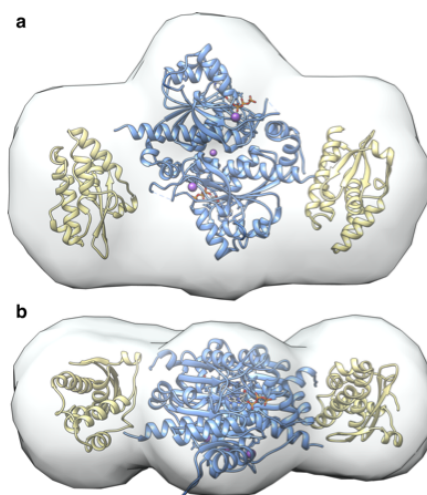


Figura 4.38. Modelo generado mediante SAXS del complejo TcdA-CsdE formado mediante BMOE. Forma DAMMIF (Franke and Svergun, 2009) representativa (superficie blanca), reconstruida mediante la imposición de simetría P2 con TcdA (azul) y CsdE (amarillo) como modelos. Se muestran dos orientaciones con diferencia de 90° de rotación alrededor del eje horizontal (a y b).

4.8. Análisis de la interacción de TcdA con ARNt^{Lys}(UUU)

4.8.1. Ensayos de EMSA (unión de TcdA con ARNt)

Con el fin de demostrar la interacción entre TcdA y el ARNt^{Lys}(UUU) se llevaron a cabo ensayos de tipo EMSA (*Electrophoretic Mobility Shift Assay*). Para ello se incubaron diferentes concentraciones de TcdA purificada (3 y 15 μ M) con 50 μ M de ARNt^{Lys}(UUU), purificado (30 minutos 293 K) en presencia de 1 mM de ATP-Mg²⁺ o de AMP-Mg²⁺. Del mismo modo se llevaron a cabo controles con ARNt solo y proteína sola.

Las muestras fueron analizadas en un gel de poliacrilamida al 8% en condiciones nativas, empleando para ello 0,5X TBE. Se mezclaron 15 μ L de la muestra más 5 μ L de *NativePAGE Sample Buffer* (50 mM Tris-HCl pH 7,6, 5% glicerol, 5 mM EDTA pH 8,0 y 0,02% *bromophenolblue*). El gel se corrió a 277 K, 12 mA durante 1 hora. Una vez finalizado éste fue teñido primeramente con *Coomassie Brilliant Blue* y después con *Toluidine Blue*. En la figura 4.39 se muestra el gel resultado del ensayo de EMSA.

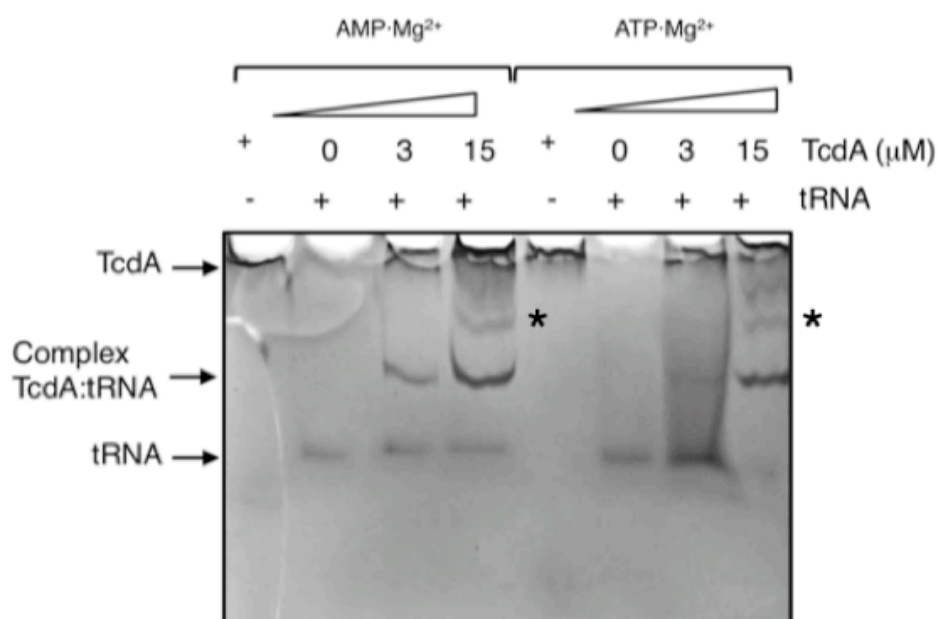


Figura 4.39. Análisis mediante EMSA de la interacción TcdA-ARNt^{Lys}(UUU). La proteína fue previamente acomplejada con AMP o con ATP para estudiar su influencia en la unión con la molécula de ARN. El asterisco muestra un posible estado del complejo proteína-ARN de mayor peso al formado por (TcdA-ARNt)₂. Por su parte, la proteína no acomplejada es incapaz de penetrar en el gel de acrilamida debido a su pl.

Se observa cómo la proteína acomplejada con la molécula de ARN es capaz de penetrar en el gel. El ARN solo o no unido a la proteína corre de forma muy diferente al complejo. Los asteriscos muestran complejos de mayor peso molecular al formado por (TcdA-ARNt)₂. Por su parte la proteína por sí sola es incapaz de introducirse en el gel, hecho explicado por su alta basicidad (pI 8,8), por lo que se puede observar concentrada en el pocillo (zona superior del gel).

4.8.2. Ultracentrifugación analítica del complejo TcdA-ARNt^{Lys}(UUU)

Con el fin de estudiar la unión entre TcdA y el ARNt^{Lys}(UUU), se llevaron a cabo experimentos de ultracentrífuga analítica, tanto SV (velocidad de sedimentación) como SE (equilibrio de sedimentación), en una ultracentrífuga analítica Beckman Coulter Proteomelab XL-I, la cual permitió llevar a cabo los experimentos a 293 K.

Todos los componentes ensayados, TcdA, ARNt y complejo TcdA-ARNt, se encontraban en el tampón 20 mM HEPES pH 7,4, 100 mM KCl, 50 mM NaCl, procedimiento llevado a cabo mediante filtración en gel en la columna Superdex200 10/300GL previamente equilibrada con este tampón.

Los experimentos de SV se llevaron a cabo a 20.000 rpm, monitorizando la absorbancia a 290 nm en intervalos de 1 minuto. Las distribuciones de los coeficientes de sedimentación se calcularon mediante la distribución continua $c(s)$ a través del modelo de la ecuación de Lamm. Los valores experimentales de s (sedimentación) se corrigieron bajo condiciones estándar en SEDNTERP (Laue *et al.*, 1992) para derivar los valores estándar s ($s_{20, w}$).

Los experimentos de SE se llevaron a cabo a velocidades entre 8.000-12.000 rpm, monitorizando las muestras a 280 y 290 nm. Tras el último paso de equilibrio, se llevó a cabo una carrera a gran velocidad (48.000 rpm), con el fin de estimar la línea base. Las medias de los pesos moleculares de flotación de las especies surgidas de la interacción entre la proteína y el ácido nucleico fueron determinadas mediante la validación de los datos experimentales empleando el programa HeteroAnalysis (Cole 2004; Ucci y Cole, 2004).

Con estos experimentos se pudo caracterizar tanto el tamaño como la estabilidad del complejo formado entre la proteína TcdA y el ARNt, donde siempre se añadió exceso de ARNt para mantener el complejo el mayor tiempo posible formado. Los perfiles de SV-AUC obtenidos son consistentes con la formación del complejo proteína-ARNt, el cual posee un coeficiente de sedimentación de 8,2 S. Los componentes aislados

Resultados

dieron coeficientes de sedimentación menores, siendo de 3,8 S para el dímero de TcdA y de 2,4 S para el ARNt.

En cuanto a los experimentos de SE-AUC, las medidas de la masa de flotación para el complejo dan un resultado de 114.297 ± 824 Da, encajando con un complejo formado por el dímero de TcdA más dos moléculas de ARNt ((TcdA-ARNt)₂). La masa de flotación obtenidas para el dímero de TcdA fue de 63.531 ± 840 Da y para el ARNt de 26.066 ± 156 Da. En la figura 4.40 se representan los resultados obtenidos para los ensayos de SV y SE.

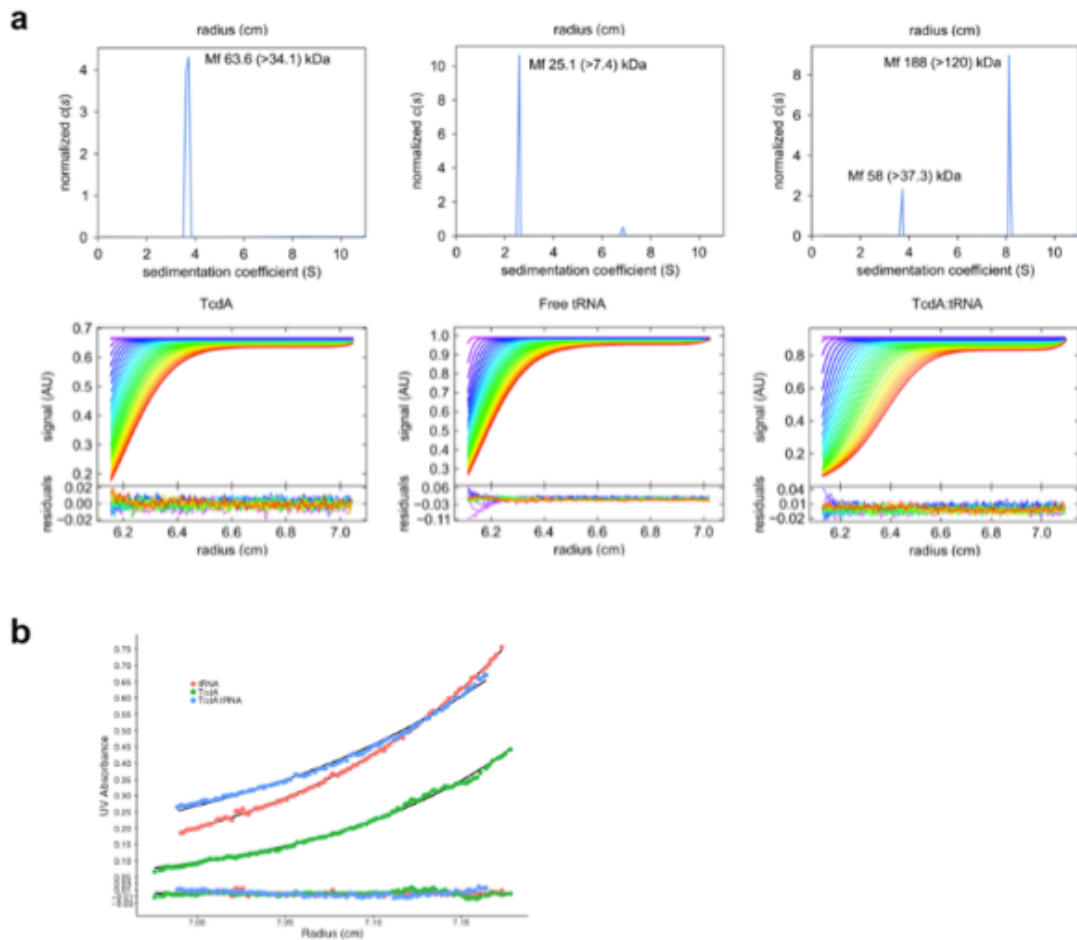


Figura 4.40. Análisis del complejo TcdA-tRNA^{Lys}(UUU) mediante AUC. **a** Representación de los resultados del experimento de SV-AUC afirmando la formación de un complejo entre TcdA y el ARNt. La distribución continua de los coeficientes de sedimentación (S) es representada para el dímero de TcdA, para el ARNt y para el complejo formado entre estos dos. Las medidas de sedimentación representativas y los residuales después del modelaje se muestran en esta figura. Para cada distribución $c(S)$ se muestran la masa molecular estimada y la menor masa compatible con el experimento. **b** Las absorbancia a 280 nm se presenta frente a la longitud de la columna para SE-AUC en los tres casos ensayados, proteína, ARN y complejo. Los residuales después del modelaje también son mostrados.

4.8.3. SAXS del complejo TcdA- ARNt^{Lys}(UUU)

Gracias a los ensayos de SAXS realizados sobre el complejo formado por TcdA-tRNA^{Lys}(UUU) fue posible generar la forma del complejo, añadiendo exceso del ARNt. De esta forma se incubaron los componentes (40 μ M de proteína y 60 μ M de ARNt) en el tampón 20 mM sodio/potasio de fosfato pH 7,4, 300 mM NaCl, 2 mM β -mercaptoetanol, para a continuación llevar a cabo una filtración en gel acoplada a SAXS, y realizar el experimento en línea. De esta forma fue posible coleccionar información de dispersión tanto del complejo como de la molécula de ARNt por separado.

Los datos obtenidos tanto del complejo como del ARNt fueron procesados e incorporados conjuntamente en base a un radio de giro (R_g) experimental coherente, dando un R_g para el complejo de 33,4 Å y para el ARNt de 23,0 Å. Así fue posible el cálculo de la masa molecular para el complejo, estimada en 90,0 kDa, coherente con la existencia de un dímero de TcdA unido a dos moléculas de ARNt. También fueron coleccionados datos mediante SAXS del dímero de TcdA (40 μ M), validando la estructura resuelta a partir de los cristales de esta proteína.

La restauración de la forma *ab initio* a partir de los datos de SAXS permitió generar una envoltura molecular para el complejo y el ARNt libre mediante el empleo de DAMMIF (Franke y Svergun, 2009) y MONSA (Svergun, D., 1999; Svergun DI y Nierhaus KH, 2000). MONSA permite el modelaje para los complejos con diferentes contrastes. Sin embargo, requiere que los componentes estructurales no sufran cambios conformacionales. Análisis preliminares de las formas restauradas sugieren que los cambios conformacionales en TcdA son probables en cuanto a su necesidad de unir la molécula de ARNt, por lo que se usó DAMMIF (Franke y Svergun, 2009) para los siguientes análisis.

El complejo formado por TcdA-ARNt posee una simetría doble, por lo que se generaron modelos de cuentas imponiendo tanto una simetría P1 como P2. Las formas obtenidas con P2 fueron capaces de reproducir mejor la simetría conocida a partir del dímero de TcdA. Así se obtuvo la forma completa del complejo obtenida mediante SAXS, el cual contiene un cuerpo central, correspondiente al dímero de la proteína, que ocupa el 75% del volumen. Esta zona central posee dos abultamientos en los laterales posicionados de forma simétrica respecto al eje, identificándolas como las moléculas de ARNt (figura 4.41).

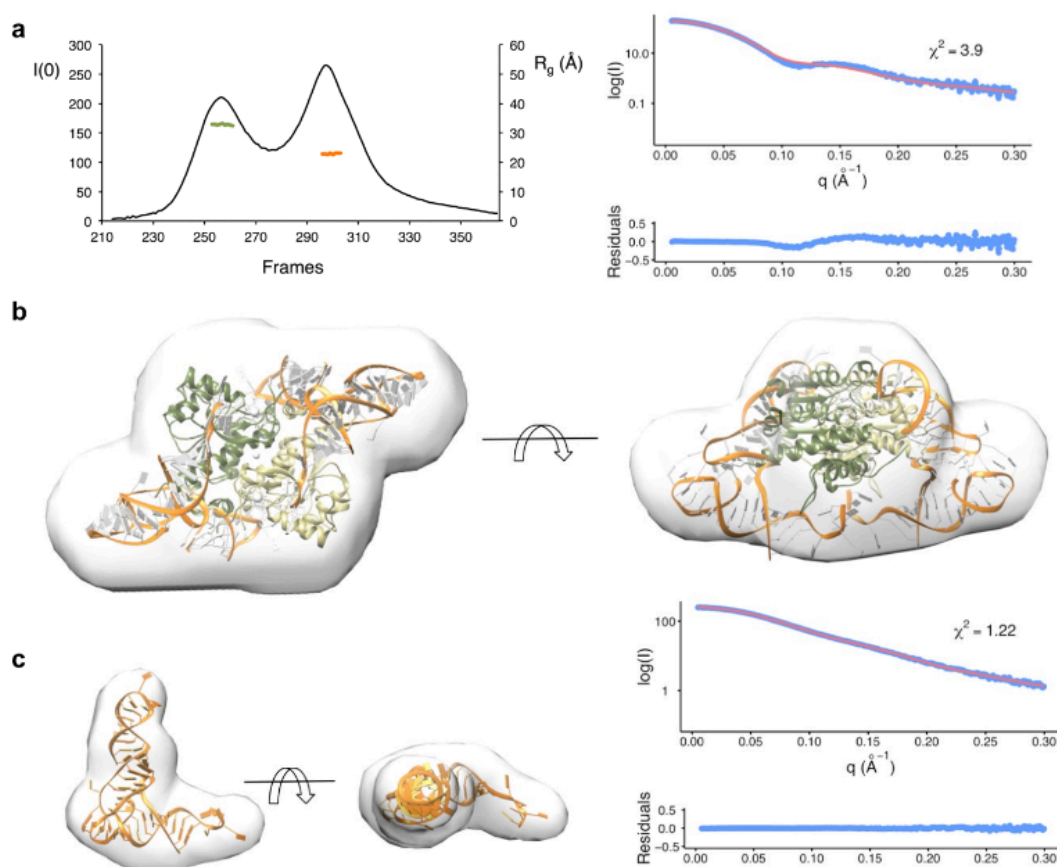


Figura 4.41. Modelo estructural del complejo TcdA-ARNt^{Lys}(UUU) determinado mediante SAXS. **a** Datos de SAXS midiendo tanto el complejo como el ARNt libre. En la gráfica se representan la intensidad a ángulo cero [$I(0)$] (eje izquierdo) y el radio de giro (R_g) (eje derecho), ambos frente a la colección de datos. En las zonas comprendidas entre 253-263 (eje X), se representa en verde la forma de rotación del complejo TcdA-ARNt y en las zonas 304-310 el ARNt, en rojo. **b** El mejor modelo calculado para el complejo revestido por la forma *ab initio* de SAXS calculada mediante DAMMIF (Franke y Svergun, 2009). La estructura cristalina de TcdA-ATP se representa en verde, mientras que el ARNt se representa en dorado. El modelo (línea roja) para los datos experimentales (puntos azules) fueron calculados con CRY SOL (Svergun DI, 1995). χ^2 y los residuales también se muestran. **c** Como en el apartado (**b**), pero representando la molécula de ARNt. En este caso, el modelo es un cuerpo rígido del modelo de ARNt dentro de la envoltura de *ab initio* de SAXS.

Con estos datos se llevó a cabo el modelaje del complejo dentro de la envoltura generada por SAXS. Como ayuda se empleó la estructura cristalográficamente resuelta de TcdA-ATP y una molécula representativa de ARNt (pdb 4jxx). La forma y el

Resultados

volumen de la envoltura de SAXS indican que TcdA ocupa la parte central, con dos moléculas de ARNt unidas simétricamente a sus lados. Este hecho involucra a numerosos residuos cargados positivamente de la proteína, pudiendo interactuar con las cargas negativas de los fosfatos del ARNt.

4.9. PRPP sintetasa (PrsA) de *E. coli*

4.9.1. Clonaje de PrsA

El gen codificante para la PRPP sintetasa se obtuvo del genoma de *E. coli* (XL1blue) mediante la técnica PCR con oligonucleótidos específicos para este gen (anexo I, tabla 1). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes se clonó en el vector de expresión con resistencia a kanamicina pETM11. En este caso el vector y el inserto no fueron digeridos con las mismas enzimas, siendo el primero digerido con las enzimas de restricción *NcoI/BamHI* y el segundo con *BsaI/BamHI*. Una vez obtenida la construcción pETM11-(PRPP sintetasa), se transformó en la cepa de clonaje de *E. coli* XL1blue, creciendo colonias en placas de LB con agar al 1,5% con kanamicina 50 µg/mL. Una vez comprobado que la construcción era correcta mediante PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen (anexo I, tabla 1), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con kanamicina.

4.9.2. Expresión de PrsA

Posteriormente se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Esta prueba se realizó en 1 mL de LB más kanamicina a la concentración mencionada anteriormente. Las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación se disminuye la temperatura hasta 293 K y se induce el cultivo con la adición de 0,5 mM IPTG (3 horas a 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonificado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 µL del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más kanamicina en matraces de 2 L de capacidad siguiendo el mismo procedimiento.

Sin embargo, el tiempo de expresión se llevó a 12 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

4.9.3. Purificación de PrsA

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular fue resuspendido en el tampón de lisis (50 mM H_2KPO_4 , 50 mM Tris-HCl, 300mM NaCl, 25 mM imidazol pH 7,5), al que se le añade 10 mM de PMSF. Una vez resuspendido el sedimento celular, éste es sonificado (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. A continuación el lisado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 μm de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de una ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando la concentración de imidazol a 250 mM en un solo paso donde se recogieron las alícuotas de elución (figura 4.42). Las muestras fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 35 kDa, correspondiente a la masa molecular del monómero de la PrsA (figura 4.42).

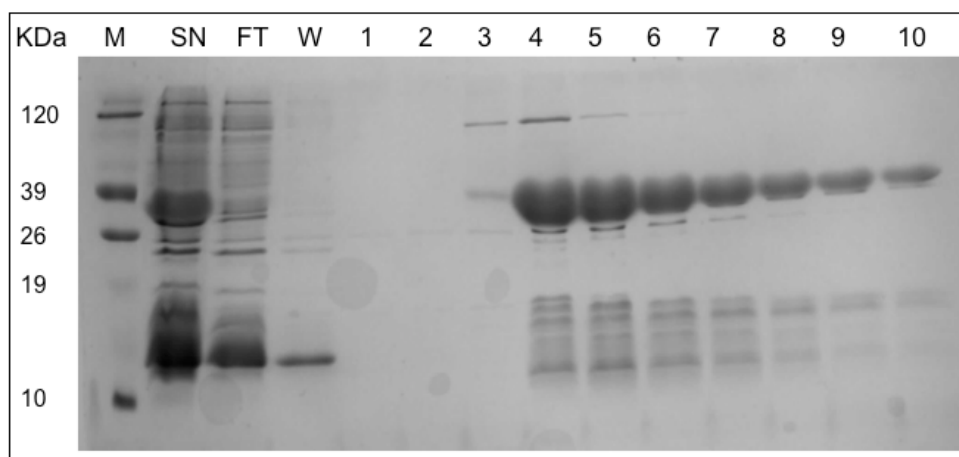


Figura 4.42. Purificación mediante IMAC de la PrsA. Mediante el análisis por PAGE-SDS se muestran las alícuotas de la purificación, observándose que la mayoría de la proteína, se encuentra sobre 35 kDa, corresponde a la recombinante PrsA. M es el marcador, SN es la muestra tras la lisis celular y el filtrado, FT es el *flow-through*, W es el lavado y los números corresponden a las alícuotas analizadas.

Resultados

Las alícuotas obtenidas del paso anterior fueron concentradas mediante el empleo de concentradores (*Millipore*) y cargadas en una columna de filtración en gel HiLoad 16/60 Superdex200, previamente equilibrada con el tampón 20 mM H₂KPO₄, 20 mM Tris-HCl, 150 mM NaCl pH 7,5. La filtración en gel fue llevada a cabo a 1 mL/min, generando un pico con volumen de elución de 77,59 mL, correspondiente en este caso a un homohexámero (figura 4.43). Las muestras obtenidas en el pico fueron analizadas mediante PAGE-SDS, donde se puede observar la gran pureza de la proteína obtenida.

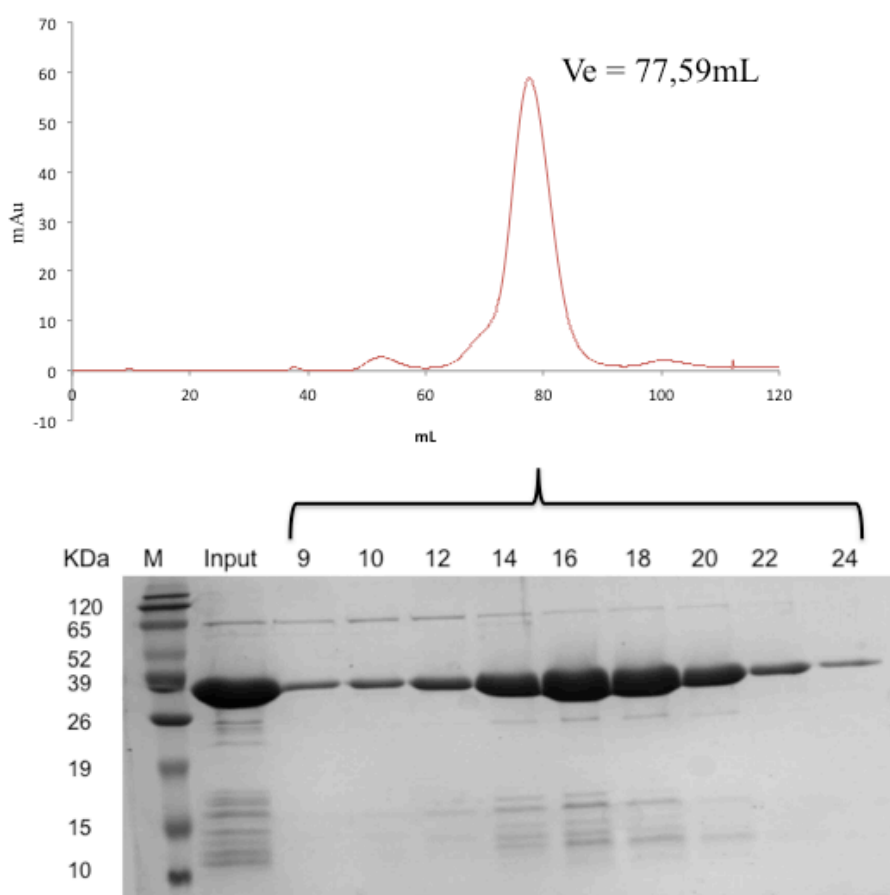


Figura 4.43. Purificación mediante filtración en gel de PrsA. Se muestra el cromatograma generado, representando las miliunidades de absorbancia (mAU) a 280 nm frente al volumen. Se puede observar que el volumen de elución (Ve) es de 77,59 mL correspondiente al estado de homohexámero de la proteína. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. La etiqueta M es el marcador, *Input* es la muestra cargada y los números corresponden a las alícuotas analizadas.

4.9.4. Cristalización de PrsA en complejo con ATP

La PrsA fue cristalizada mediante el método de gota sentada (difusión de vapor) a 293 K, acompañada con ATP, empleando proteína purificada a 6,5 mg/mL (previamente analizada por DLS) y ATP a 10 mM. Se añadió 10 mM de MgCl₂. La solución de cristalización empleada en este caso fue 0,2 M K₂HPO₄, 20% (v/v) PEG 3.350 pH 9,9. Los cristales aparecieron a los 4 días, siendo congelados en nitrógeno líquido al 20% de glicerol.

4.9.5. Colección y procesamiento de datos de difracción de los cristales de PrsA acompañados con ATP

Los cristales fueron difractados en el haz de luz ID14-4 en el sincrotrón ESRF (Grenoble, Francia). Los cristales difractaron a 1,77 Å y pertenecen al grupo espacial P6₅22 con una celda unidad de a=105,64 Å, b=105,64 Å, c=314,58 Å con los ángulos $\alpha=\beta=90$ y $\gamma=120$ (tabla 4.9). En la figura 4.44 se muestra la imagen de difracción del cristal acompañado con ATP.

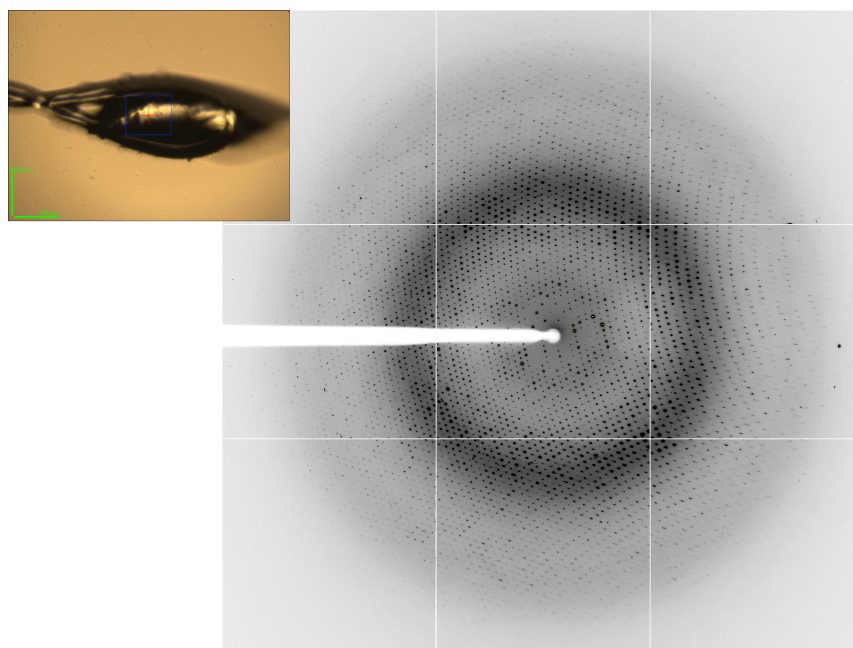


Figura 4.44. Imagen de difracción de los cristales de *EcPrsA* en presencia de ATP. En el lado izquierdo superior se puede observar la imagen del cristal en el *loop* listo para difractar. A la derecha se muestra la imagen de difracción obtenida a 1,77 Å de resolución.

4.9.6. Procesamiento de los datos, resolución estructural y refinamiento de PrsA en complejo con ATP

Los datos de la estructura fueron integrados con XDS (Kabsch, 2010) y escalados con Aimless (Evans y Murshudov, 2013). Para la solución estructural se utilizó el reemplazo molecular con el programa PHASER integrado en la suite de PHENIX (McCoy *et al.*, 2007) usando como modelo la estructura de un monómero de PrsA de *Burkholderia pseudomallei* (pdb 3dah) (Baugh *et al.*, 2013; Henriksen *et al.*, 1996). Se localizaron tres cadenas de EcPrsA en la unidad asimétrica correspondiente a medio hexámero. La otra mitad del hexámero se genera gracias a un eje cristalino binario.

El trazado de la estructura de la proteína, comprobación de las cadenas laterales, y moléculas de solvente fue realizada en ciclos que incluían la comprobación visual en el ordenador mediante el programa gráfico Coot (Emsley *et al.*, 2010) y el refinamiento de la estructura con phenix.refine (Afonine *et al.*, 2010). Finalmente la estructura fue validada con MolProbity (Chen *et al.*, 2010). En la tabla 4.9 se muestran las estadísticas obtenidas de estos procesos.

<i>EcPrsA</i> -ATP	
pdb code	-
Data collection	
Wavelength (Å)	0.9322
Resolution range (Å)	45.49-1.77 (1.83-1.77)
Space group	$P 6_5 2 2$
a, b, c (Å)	105.64, 105.64, 314.58
β (°), $\alpha = \gamma = 90^\circ$	90, 90, 120
Total reflections	774,688 (75,134)
Unique reflections	101,067 (9836)
Multiplicity	7.7 (7.6)
Completeness (%)	99.18 (98.36)
Mean I/ σ (I)	14.45 (2.77)
Wilson B-factor	22.47
R-merge	0.103 (0.979)
R-meas ^a	0.111 (1.052)
CC1/2 ^b	0.040 (0.379)
CC* ^c	0.999 (0.907)
Refinement	
R-work	0.1861 (0.3044)
R-free	0.2204 (0.3545)
# non-H atoms	7648
# Protein	7026
# Ligand atoms	129
# Water	493
Protein residues	911
RMS (bond) (Å)	0.016
RMS(angles)(°)	1.53
Ramachandran analysis	
Favored/Allowed/Outlier (%)	95.2/4.5/0.3
Clashscore	7.65
Average B-factor (Å ²)	35.57
Protein	34.56
Ligands	66.06
Solvent	42.06

Tabla 4.9. Datos estadísticos de la solución de la estructura de *EcPrsA* acomplexada con ATP.

4.9.7. La estructura cristalográfica de *EcPrsA* en complejo con ATP

La estructura cristalográfica de *EcPrsA* ha sido determinada en presencia de ATP. De forma independiente, una estructura de *EcPrsA* en forma apo (en presencia de Mg^{2+}) ha sido resuelta recientemente a 2,7 Å (pdb 4s2u, 26 junio 2016). *EcPrsA* pertenece a la clase I de PRPP sintetasas, igual que las de *Bacillus subtilis* (Eriksen *et al.*, 2002),

Resultados

Burkholderia pseudomallei (pdb 3dah) (Baugh *et al.*, 2013) y la humana (Li *et al.*, 2007), cuyas estructuras también se conocen. Las PRPP sintetasas de clase I son moléculas con dos dominios en cada monómero de plegamientos similares y con un alto grado de identidad de secuencia. Adoptan estructuras cuaternarias hexaméricas, donde los seis protómeros que constituyen los hexámeros se organizan en un trímero de dímeros relacionados en un eje de simetría ternario. Estas estructuras hexaméricas se han descrito como hélices donde los dominios N-terminales se orientan hacia la zona central del hexámero, mientras que los dominios C-terminales se giran hacia el exterior del hexámero (figura 4.45).

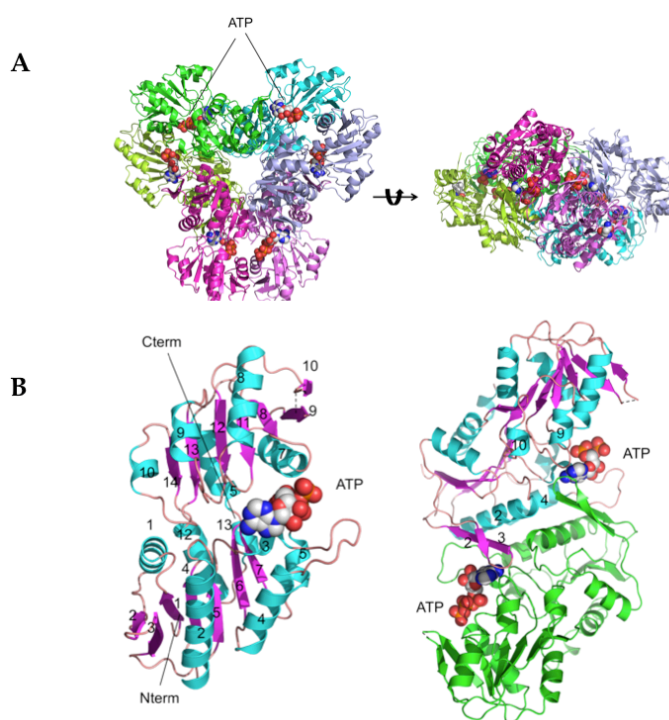


Figura 4.45. Estructura tridimensional de *EcPrsA* formando complejo con ATP. A) Representación en dos orientaciones del complejo de *EcPrsA* con ATP. El ligando se halla representado en modelo de bolas. La orientación de ambas figuras, una respecto de la otra, es de 90°. **B)** Representación de la estructura del monómero.

Los dos dominios de la *EcPrsA* son notablemente similares en estructura y, en mucha menor medida, en secuencia (la identidad de secuencia es < 20%). El dominio N-terminal (29-173) y el C-terminal (188-317) pueden superponerse con un r.m.s.d. de 2,5 Å a lo largo de 108 átomos C α (figura 4.46). Las mayores discrepancias ocurren a nivel

de inserciones/delecciones, lazos de longitud variable, y en aquellas regiones especializadas en formar las interfaces de interacción entre los distintos monómeros. Cada dominio tiene una estructura α/β abierta con una lámina β central compuesta por 5 hebras β rodeada por 3 hélices α en una cara y una hélice α en la otra cara, además de encontrarse una hélice 3_{10} en el dominio N-terminal, que está ausente en el C-terminal. Como se puede observar en la figura 4.45.b), ambos dominios están flanqueados por una larga lámina β antiparalela (hebras β 1 y 2 en el dominio N-terminal y 9 y 10 en el C-terminal) que empaqueta contra el flanco de la lámina β central (1 y 2) y dos hélices α (7 y 8).

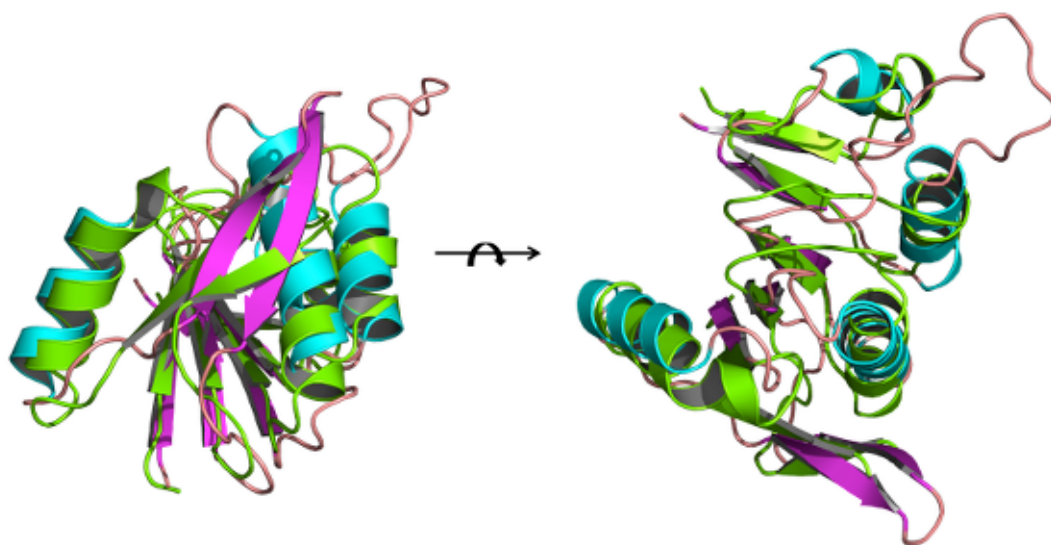


Figura 4.46. Superposición de los dominios N-terminal y C-terminal de *EcPrsA*. El dominio N-terminal está coloreado por estructura secundaria (ciano, hélices α ; violeta, hebras β ; salmón, lazos) y el dominio C-terminal en color verde.

El sitio activo de PrsA (figura 4.47) en la estructura cristalográfica está ocupado por una molécula del sustrato ATP, que permite definir con exactitud los residuos del sitio activo que participan en la estabilización del ATP. La cavidad donde se une el ATP la forman residuos de dos cadenas diferentes. Las cadenas laterales de Asn123A, Arg125A y His157A interaccionan con el grupo ribosa-trifosfato del ATP, mientras que las cadenas laterales de Arg122A, Phe61A, Asp63A y Glu65A participan en la unión del anillo de adenina del ATP y el mantenimiento de su posición y orientación (figura 4.47). De acuerdo con la estructura cristalina, el ATP se sitúa por tanto en el surco formado por la confluencia de dos cadenas vecinas y, como tal, justifica que el centro activo de *EcPrsA* esté compuesto por aminoácidos de ambos protómeros. Varios de los residuos que interaccionan con ATP se encuentran en zonas particularmente flexibles de la

Resultados

estructura. Así, la Phe61 se localiza en la punta de las dos hebra β que flanquean los dominios, y la Arg122 se encuentra en un lazo muy largo y flexible (residuos 119-135).

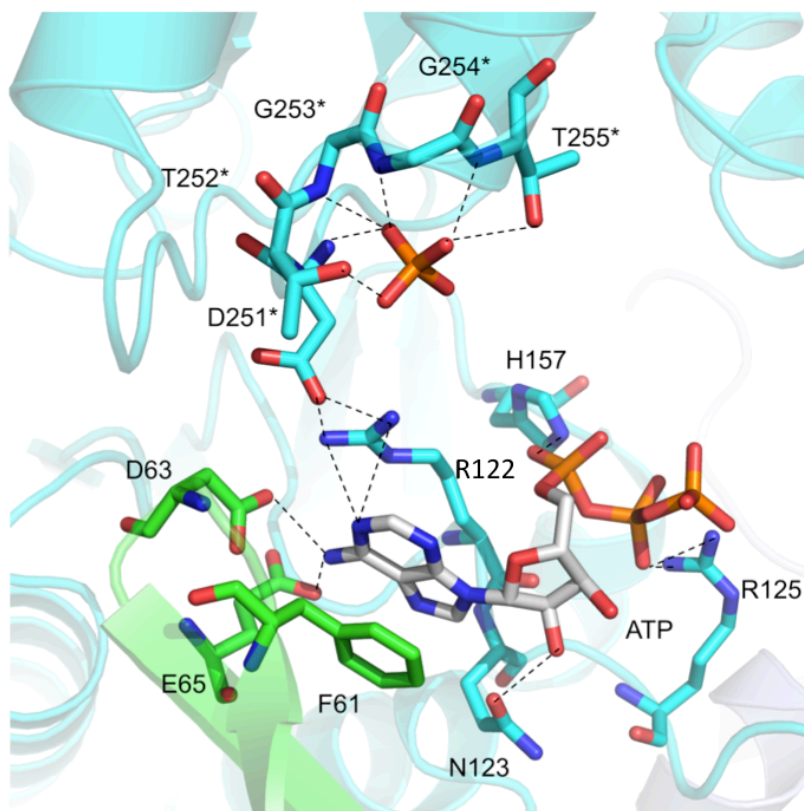


Figura 4.47. Centro activo de *EcPrsA* en presencia de ATP. El centro activo está compuesto por residuos procedentes de más de un monómero del hexámero de *EcPrsA*. En la representación ambas cadenas se muestran en diferente color (cian y verde). La molécula de fosfato se haya interaccionando con el motivo *S-loop* entre los residuos Asp251-Thr255 en el extremo N-terminal de una hélice α .

En la misma cavidad en el lado opuesto al ATP se observa un ión fosfato proveniente de la condición de cristalización, que está fuertemente asociado con un segmento de PrsA formado por los residuos Asp251B, Thr252B, Gly253B, Gly254B y Thr255B. Estos residuos, que pueden identificarse mediante alineamiento de secuencias con otras PRPP sintetasas como el motivo de unión a PRPP, establecen interacciones con el grupo fosfato mediante sus cadenas laterales (los dos residuos de Thr) o con el enlace peptídico (el resto). De hecho, las conformaciones del esqueleto polipeptídico de este segmento son inusuales y están fuertemente conservadas en muchas PRPP sintetasas. En concreto, este motivo Asp251-Thr255 es responsable de la unión de la parte ribosa 5-

fosfato (R5P) del PRPP. También mediante alineamientos de secuencia es posible inferir el sitio de unión al pirofosfato de PRPP, que se correspondería con los residuos Asp196A-Gly199A.

Es interesante notar que la cadena lateral de Asp251B también contribuye a formar el sitio de unión del ATP mediante la formación de un enlace de hidrógeno fuerte con el grupo guanidino de la cadena lateral de Arg122A (figura 4.47). El requerimiento de residuos de la cadena vecina para la estabilización del ATP genera el sitio activo compuesto y permite vincular funcionalmente las dos subunidades.

4.10. *EcOPRT* y *MtOPRT*

4.10.1. Clonaje de *EcOPRT*

El gen codificante para la OPRT de *E. coli* se obtuvo del genoma de esta misma bacteria (XL1blue) mediante la técnica PCR y el uso de los oligonucleótidos específicos para este gen (anexo I, tabla 1). Una vez obtenido el fragmento de PCR y digerido correctamente con las enzimas de restricción correspondientes se clonó en el vector de expresión con resistencia a ampicilina pET21b. El oligonucleótido 3' empleado en este caso contiene un codón *stop* con el fin de no añadir la cola de histidinas C-terminal que codifica el vector. Tanto el vector como el inserto fueron digeridos con las enzimas *NdeI*/*XhoI*. Una vez obtenida la construcción pET21b-*EcOPRT*, se transformó en la cepa de clonaje de *E. coli* XL1blue, creciendo colonias en placas de LB con agar al 1,5% con ampicilina 100 µg/mL. Tras comprobar que la construcción era correcta mediante PCR de colonia y secuenciación empleando los oligonucleótidos para el promotor T7 y el 3' específico de gen (anexo I, tabla 1), se transformó con el ADN obtenido de minipreparaciones en la cepa de expresión BL21(DE3), plaqueando de nuevo en LB-agar con ampicilina.

4.10.2. Expresión de *EcOPRT*

Posteriormente se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Estas pruebas se realizaron en 1 mL de LB más ampicilina a la concentración mencionada anteriormente. Las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación se mantiene la temperatura y se induce el cultivo con la adición de 1 mM IPTG (3 horas 800 rpm). Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonificado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000

Resultados

g (10 minutos 277 K), para tomar 15 μ L del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de LB más ampicilina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin embargo, el tiempo de expresión se llevó a 5 horas y el sedimento celular obtenido se guarda a 195 K hasta ser procesado.

4.10.3. Purificación de *EcOPRT*

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular fue resuspendido en el tampón de lisis (25 mM Tris-HCl pH 7,6, 0,5 mM EDTA y 2 mM β -mercaptoetanol), al que se le añade 10 mM de PMSF. La muestra fue sometida a un golpe de calor (10 minutos 340 K). A continuación ésta fue sonicada (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. El lisado generado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante es filtrado empleando filtros de 0,45 μ m de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de intercambio aniónico, equilibrada previamente con el tampón de lisis mediante el uso de una ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando gradualmente la concentración de NaCl hasta 1.000 mM, sin cambiar ninguno de los otros componentes del tampón. La proteína eluyó a lo largo del gradiente (figura 4.48). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 25 kDa, correspondiente a la masa molecular del monómero de *EcOPRT* (figura 4.48).

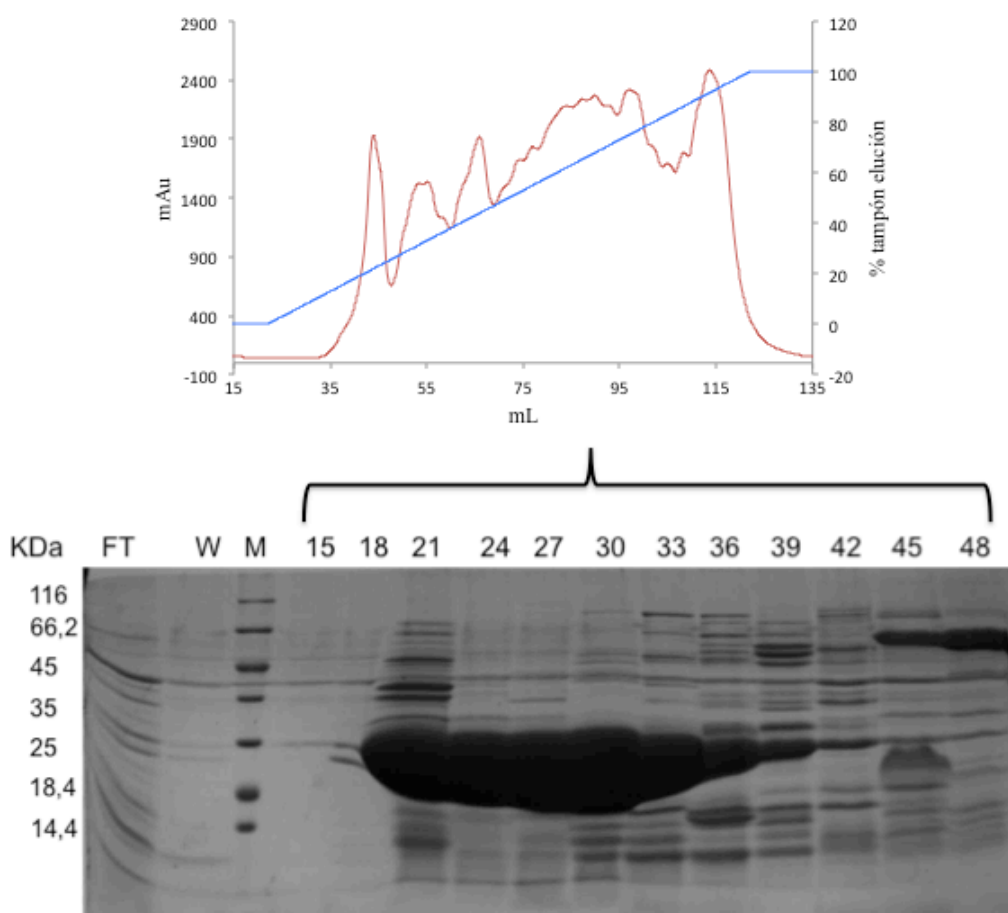


Figura 4.48. Purificación mediante intercambio iónico de EcOPRT. En el cromatograma se representan las miliunidades de absorbancia (mUA) a 280 nm (línea roja) frente al volumen de elución. La línea azul representa la concentración del tampón de elución, observándose el gradiente. En la PAGE-SDS se muestran las muestras de la purificación, observándose que la mayoría de la proteína, sobre 25 kDa, corresponde a la recombinante EcOPRT. Se puede observar la gran pureza de la muestra. En el gel de PAGE-SDS M es el marcador, FT el *flow-through*, W el lavado y los números corresponden a las alícuotas analizadas.

Las alícuotas analizadas en el paso anterior mediante PAGE-SDS fueron concentradas mediante el uso de concentradores (*Millipore*), para ser cargadas en una columna HiLoad 16/60 Superdex200 previamente equilibrada en el tampón de lisis. La filtración en gel se llevó a cabo a un flujo de 1 mL/min. La muestra dio un pico de elución de 93,84 mL, correspondiente al homodímero de la proteína (figura 4.49). Las muestras de elución fueron analizadas mediante PAGE-SDS.

Resultados

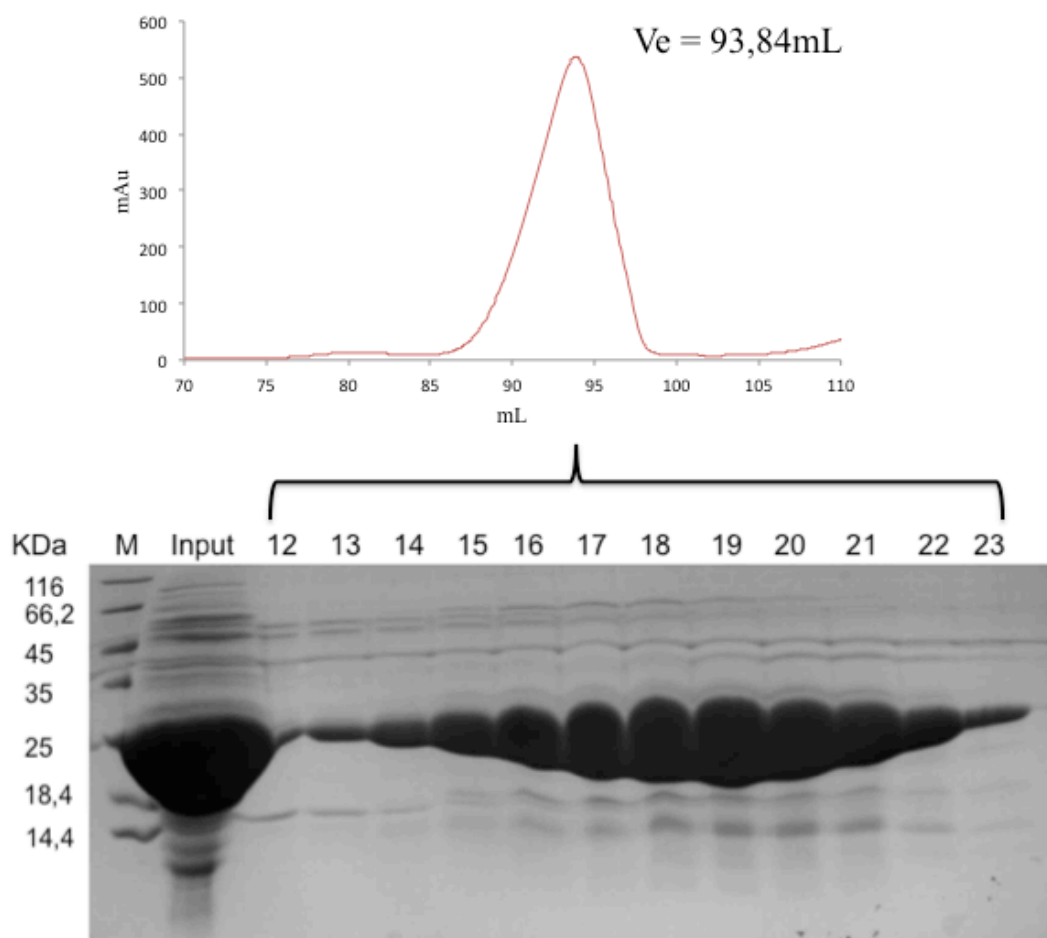


Figura 4.49. Purificación mediante filtración en gel de *EcOPRT*. Se muestra el cromatograma generado, representando las miliunidades de absorbancia (mUA) a 280 nm frente al volumen. Se puede observar que el volumen de elución (V_e) es de 93,84 mL correspondiente al estado oligomérico de homodímero de la proteína. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M es el marcador, *Input* es la muestra cargada en la columna y los números corresponden a las alícuotas analizadas.

4.10.4. Clonaje de *MtOPRT*

El gel codificante para la OPRT de *Mycobacterium tuberculosis* fue insertado en el vector de expresión pETM11. En esta ocasión el gen se obtuvo por síntesis, por lo que el paso de la PCR no fue necesario (GenScript). Se transformó primeramente en la cepa *XL1blue* y posteriormente, tras la comprobación de la secuencia de la construcción (oligonucleótidos T7 promotor y terminador, anexo I, tabla 1), en BL21(DE3). Los pasos

seguidos fueron los mismos que en los casos anteriores, empleando placas de LB con 1,5% de agar más kanamicina (50 µg/mL) en el caso de los plaqu coastos.

4.10.5. Expresión de MtOPRT

Posteriormente se seleccionaron dos colonias de BL21(DE3) con el fin de llevar a cabo las pruebas de expresión. Ésta prueba se realizaron en 1 mL de *Superior Broth* (AthenaES®) más kanamicina a 50 µg/mL. Las colonias fueron crecidas a 310 K hasta alcanzar una D.O.₅₉₀ de 0,6-0,8. A continuación se mantiene la temperatura y se induce el cultivo con la adición de 1 mM IPTG durante 3 horas a 800 rpm. Una vez transcurrido el tiempo de expresión el cultivo es centrifugado a 1.000 g (10 minutos 277 K). Tras este paso el sobrenadante generado es decantado y el sedimento celular es resuspendido en 400 µL de PBS pH 7,0, para ser posteriormente sonicado. Tras la lisis celular se tomaron 15 µL de la muestra, la cual fue posteriormente centrifugada a 10.000 g (10 minutos 277 K), para tomar 15 µL del sobrenadante. Estas dos alícuotas fueron analizadas mediante PAGE-SDS con el fin de comprobar la buena expresión de la proteína recombinante.

Una vez comprobada la calidad de la expresión se escaló el cultivo a 750 mL de *Superior Broth* (AthenaES®) más kanamicina en matraces de 2 L de capacidad siguiendo el mismo procedimiento. Sin embargo, el tiempo de expresión se llevó a 5 horas y el sedimento celular obtenido se guardó a 195 K hasta ser procesado.

4.10.6. Purificación de MtOPRT

Tras llevar a cabo el escalado del cultivo se realizó la purificación de la proteína recombinante. Para ello el sedimento celular fue resuspendido en el tampón de lisis (50 mM Tris-HCl pH 8,0, 300 mM NaCl y 20 mM imidazol), al que se le añade 10 mM de PMSF. Una vez resuspendido el sedimento celular, éste es sonicado (2,5 minutos 277 K) en ciclos de 30 segundos de sonicación y 20 segundos de reposo. A continuación, el lisado es centrifugado a 10.000 g (20 minutos 277 K), tras lo cual el sobrenadante generado es filtrado empleando filtros de 0,45 µm de diámetro de poro (PES). Tras este proceso la muestra es cargada en una columna de afinidad de níquel para llevar a cabo una IMAC, equilibrada previamente con el tampón de lisis. Para ello la columna es conectada y equilibrada en el tampón de lisis mediante el uso de una ÄKTA Prime. La muestra fluyó a una velocidad de 1 mL/min, llevándose a cabo la elución aumentando gradualmente la concentración de imidazol hasta 250 mM, sin cambiar ninguno de los otros componentes del tampón. La proteína eluyó a 250 mM de imidazol (figura 4.50). Las muestras correspondientes al pico de elución fueron analizadas mediante PAGE-SDS, obteniendo una banda mayoritaria de 22 kDa, correspondiente a la masa molecular del monómero de MtOPRT (figura 4.50).

Resultados

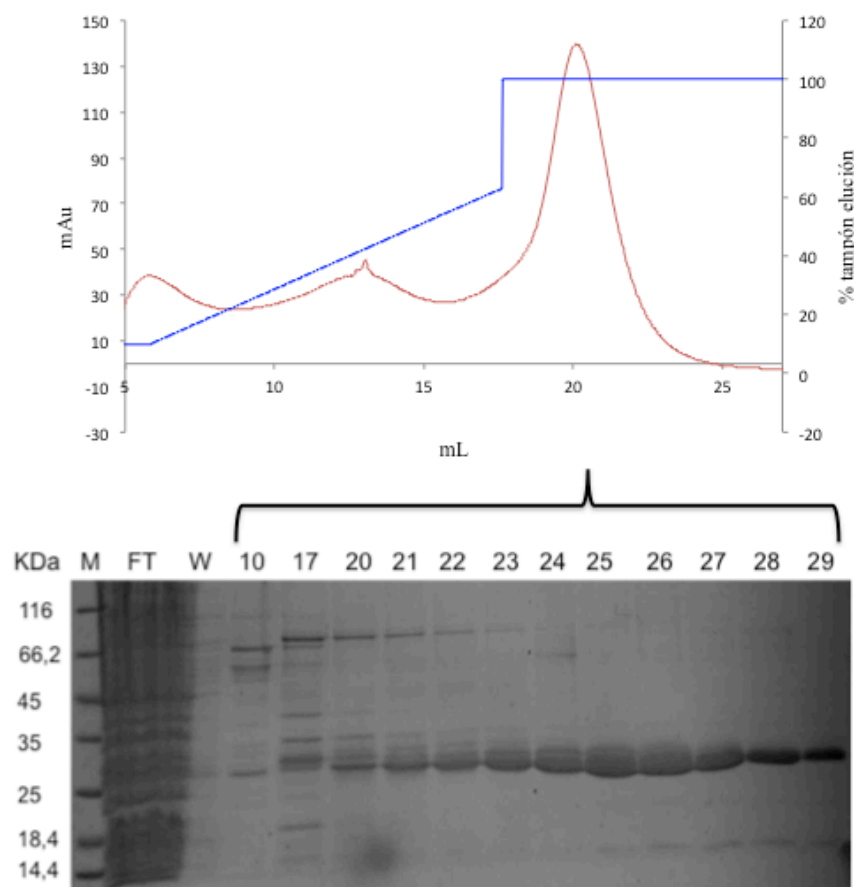


Figura 4.50. Purificación mediante IMAC de *MtOPRT*. En el cromatograma se representan las miliunidades de absorbancia (mUA) a 280 nm (línea roja) frente al volumen de elución. La línea azul representa la concentración del tampón de elución, observándose el gradiente. En la PAGE-SDS se muestran las muestras de la purificación, observándose que la mayoría de la proteína, sobre 22 kDa, corresponde a la recombinante *MtOPRT*. Se puede observar la gran pureza de la muestra. En el gel de PAGE-SDS M es el marcador, FT el *flow-through*, W el lavado y los números corresponden a las alícuotas analizadas.

Las alícuotas analizadas en el paso anterior por PAGE-SDS fueron concentradas mediante el uso de concentradores (*Millipore*), para ser cargadas en una columna Superdex75 10/300GL previamente equilibrada en 20 mM fosfato sodio/potasio pH 8,0. La filtración en gel se llevó a cabo a un flujo de 0,5 mL/min. La muestra dio un pico de elución de 10,96 mL, correspondiente al homodímero de la proteína (figura 4.51). Las muestras de elución fueron analizadas mediante PAGE-SDS, donde se puede apreciar la gran pureza de la proteína expresada.

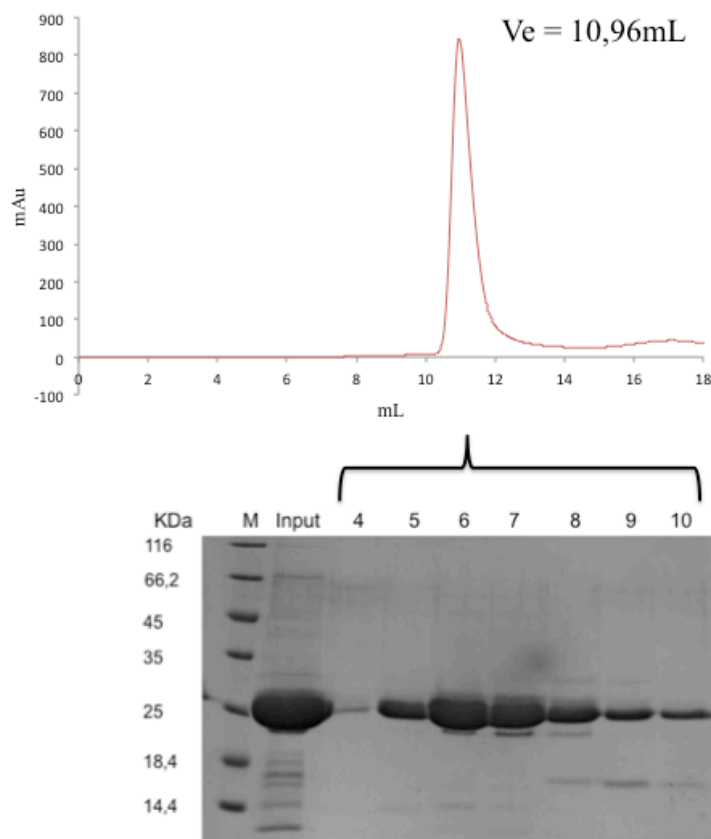


Figura 4.51. Purificación mediante filtración en gel de *MtOPRT*. Se muestra el cromatograma generado, representando las miliunidades de absorbancia (mUA) a 280 nm frente al volumen. Se puede observar que el volumen de elución (V_e) es de 10,96 mL correspondiente al estado oligomérico de un homodímero de la proteína. En el PAGE-SDS se observa la proteína con un gran grado de pureza y a la altura esperada. En éste M es el marcador, *Input* es la muestra cargada en la columna y los números corresponde a las alícuotas analizadas.

4.10.7. Cinéticas de *EcOPRT* y *MtOPRT*

4.10.7.1. Cinéticas de *EcOPRT*

NOTA: todos los experimentos de este apartado se realizaron en colaboración con Sara Gómez Quevedo, CIB-CSIC.

Una vez purificadas ambas proteínas se llevaron a cabo los ensayos cinéticos. En el caso de la *EcOPRT* se empleó un método de detección por espectrometría continua, midiendo la reacción que tiene lugar hacia la producción de OMP y PPI. Para ello se

Resultados

emplearon cubetas compatibles con la detección de luz UV (*Micro Quartz cuvette*) en un espectrofotómetro *Eppendorf BioSpectrometer*. La reacción fue llevada a cabo a 298 K durante 5 minutos, midiendo absorbancia a 295 nm. El volumen de reacción fue de 150 μL , llevándose a cabo ésta en el tampón 20 mM Tris-HCl pH 7,6. Para ello se fijó la concentración de uno de los sustratos, en este caso el PRPP (100 μM) y del MgCl_2 (2 mM), variando la concentración del otro sustrato de la reacción. De esta forma, los sustratos de los cuales se varió la concentración fueron el ácido orótico (OA) (5-100 μM), el 5-fluoroorótico (5-FOA) (5-100 μM) y 5-aminoorótico (5-AOA) (10-750 μM). Una vez mezclados todos los componentes (ácido orótico o derivado, más PRPP y MgCl_2), la reacción fue iniciada tras la adición de 300 nM de enzima purificada, empezando a medir el decrecimiento de absorbancia a 295 nm inmediatamente. Los datos obtenidos de absorbancia frente a minutos fueron convertidos a unidades cinéticas mediante el uso de los coeficientes de extinción molar (ϵ , $\text{M}^{-1} \cdot \text{cm}^{-1}$) del ácido orótico o de sus análogos y la ecuación de Lambert-Beer ($A = \epsilon \cdot c \cdot l$), donde se empleó un paso óptico (l) de 1 cm. De esta forma, los valores de ϵ a 295 nm fueron para el OA de $4.108 \text{ M}^{-1} \cdot \text{cm}^{-1}$, para el 5-FOA de $4.000 \text{ M}^{-1} \cdot \text{cm}^{-1}$ y para el 5-AOA de $3.476 \text{ M}^{-1} \cdot \text{cm}^{-1}$. Los datos obtenidos fueron ajustados a un modelo de tipo Michaelis-Menten, donde fueron calculadas las distintas constantes cinéticas gracias al programa SigmaPlot 13.0. En la tabla 4.10 se muestran los valores obtenidos, así como los de la misma enzima correspondiente al organismo *Salmonella typhimurium*. En el caso de 5-AOA como sustrato, no se detectó ninguna actividad.

Sustrato	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu \text{ M}^{-1} \cdot \text{s}^{-1}$)
OA ^a	$94,9 \pm 22,2$	26,9	0,28
5-FOA ^a	$171,8 \pm 4,2$	67,9	0,39
OA ^b	28	27,5	0,99
5-AOA ^a	-	-	-

Tabla 4.10. Constantes cinéticas de EcOPRT y StOPRT. Se muestran las constantes calculadas para la OPRT de *E. coli* (a) para OA (ácido orótico) y 5-FOA (5-fluoroorótico) y las publicadas para *S. typhimurium* (b) en Bhatia *et al.*, 1992 para OA. En el caso de 5-AOA (5-aminoorótico), no pudieron ser calculadas.

4.10.7.2. Cinéticas de *MtOPRT*

En el caso de la *MtOPRT* se empleó el mismo método descrito anteriormente, midiendo la reacción que tiene lugar hacia la producción de OMP y PPi, empleando en este caso únicamente 5-FOA como sustrato a estudiar ($5\text{-FOA} + \text{PRPP} \rightarrow \text{OMP} + \text{PPi}$). Para ello se emplearon cubetas compatibles con la detección de luz UV (*Micro Quartz cuvette*) en un espectrofotómetro *Eppendorf BioSpectrometer*. La reacción fue llevada a cabo a 298 K durante 5 minutos, midiendo absorbancia a 295 nm. El volumen de reacción fue de 400 μL , llevándose a cabo ésta en el tampón 50 mM Tris-HCl pH 8,0. Para ello se fijó la concentración de uno de los sustratos, en este caso el PRPP (500 μM) y del MgCl_2 (20 mM), variando la concentración del otro sustrato de la reacción. Se emplearon diferentes concentraciones del sustrato 5-FOA (5-100 μM). Una vez mezclados todos los componentes, la reacción era iniciada con la adición de 100 nM de enzima, empezando a medir el decrecimiento de absorbancia a 295 nm inmediatamente. Los datos obtenidos de absorbancia frente a minutos fueron convertidos a unidades cinéticas mediante el uso del coeficiente de extinción molar (ϵ) para el 5-FOA empleado en el apartado anterior ($4.000 \text{ M}^{-1} \cdot \text{cm}^{-1}$). El paso óptico (l) empleado en este caso fue el mismo del apartado anterior (1 cm). Los datos obtenidos fueron ajustados a un modelo de tipo Michaelis-Menten, donde fueron calculadas las distintas constantes cinéticas gracias al programa SigmaPlot 13.0. Las constantes se muestran en la tabla 4.11.

Sustrato	K _m (μM)	k _{cat} (s ⁻¹)	k _{cat} /K _m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
OA	10	0,61	0,061
5FOA	9,5 \pm 2,2	4,2	0,44

Tabla 4.11. Constantes cinéticas de *MtOPRT*. Se muestran las constantes calculadas para la OPRT de *M. tuberculosis* tanto con su sustrato natural ácido orótico (OA) (Breda *et al.*, 2012) como para el análogo 5-FOA (5-fluoroorótico).

4.10.8. Cristalización de *Ec*OPRT en complejo con 5-FOA y 5-AOA y *Mt*OPRT en su forma apo

*Ec*OPRT fue cristalizada mediante el empleo del método de gota sentada (difusión de vapor) a 293 K, acomplejada con 5-FOA y 5-AOA (cocrystalización), empleando proteína purificada a 8 mg/mL (previamente analizada por DLS) y 5-FOA a 5 mM o 5-AOA a 5 mM. La solución de cristalización empleada en este caso fue 2,1 M (NH₄)₂SO₄, 0,3 M citrato de sodio pH 4,8. Los cristales aparecieron a los 4 días, siendo congelados en nitrógeno líquido al 20% de glicerol.

*Mt*OPRT fue cristalizada mediante el mismo método que la enzima de *E. coli*, aunque tan solo en su forma apo. Para ello se empleó proteína a 8 mg/mL (previamente analizada por DLS) y la solución de cristalización 0,1 M acetato de sodio pH 5,0, 5% (v/v) PGA-LM y 12% (v/v) PEG 8.000. Los cristales aparecieron a los 3 días, siendo congelados en nitrógeno líquido al 15% de glicerol.

4.10.9. Colección de datos y procesamiento de datos de los cristales de *Ec*OPRT en complejo con 5-FOA y 5-AOA y *Mt*OPRT en su forma apo

La difracción de los cristales pertenecientes a la *Ec*OPRT fue colectada en la línea PROXIMA 2A (SOLEIL, París, Francia) a 1,4 Å. En cuanto a la misma enzima acomplejada con los análogos del sustrato nativo, éstos fueron difractados en la misma línea, obteniendo una resolución 1,43 Å para el complejo con 5-FOA, y de 1,62 Å para el complejo con 5-AOA. En la figura 4.52 se muestra una de las imágenes de difracción de *Ec*OPRT acomplejada con 5-FOA y 5-AOA.

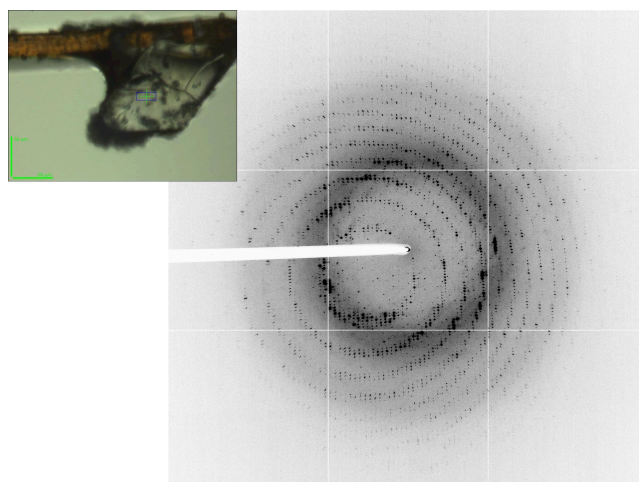
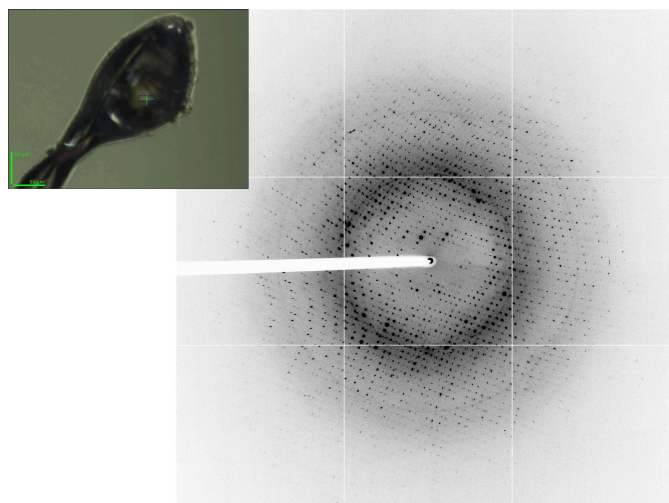
A**B**

Figura 4.52. Imagen de difracción de los cristales de *EcOPRT*. Se muestran los cristales y las imágenes de difracción de *EcOPRT* acompañada con los análogos de sustrato 5-FOA (**A**) y 5-AOA (**B**).

La difracción de los cristales pertenecientes a la *MtOPRT* fueron colectados en la línea PROXIMA 2A (SOLEIL, París, Francia) a 1,54 Å. En la figura 4.53 se muestra una de las imágenes de difracción de *MtOPRT*.

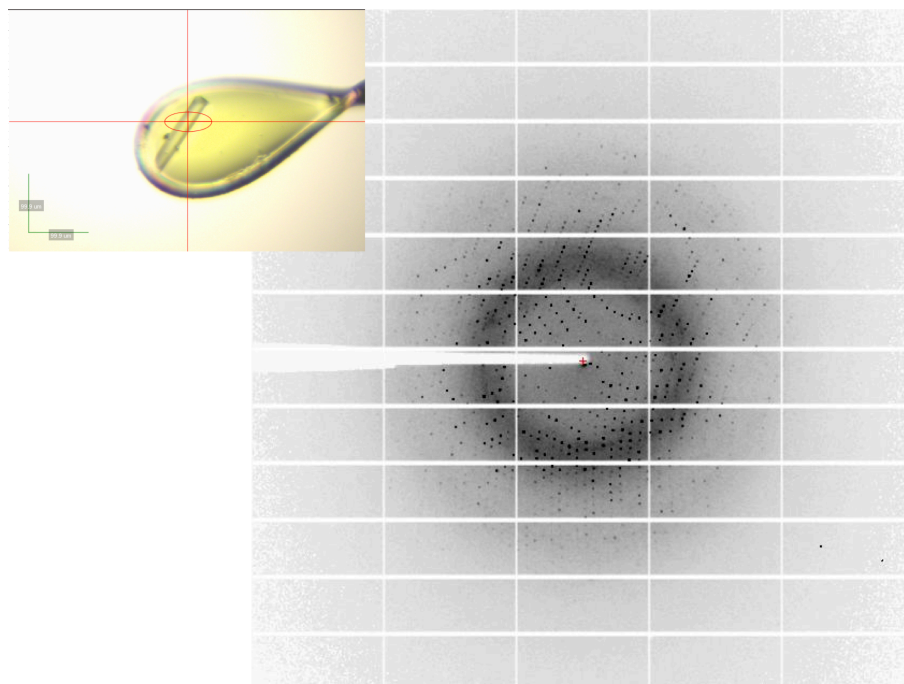


Figura 4.53. Imagen de difracción de los cristales de *MtOPRT*. Se muestran uno de los cristales difractados y la imagen de difracción de *MtOPRT* en su forma *apo*.

4.10.10. Resolución estructural y refinamiento de *EcOPRT* en complejo con 5-FOA y 5-AOA

Todos los datos de la estructura *apo* fueron integrados con XDS (Kabsch, 2010) y escalados gracias a *Aimless* (Evans y Murshudov, 2013). En cuanto a los datos obtenidos a la proteína acomplexada con los análogos, estos fueron empleados con el fin de resolver la estructura mediante la técnica de MR, gracias a la estructura apo de *EcOPRT* (pdb 1oro) (Henriksen *et al.*, 1996) con la ayuda de PHASER (McCoy *et al.*, 2007). La estructura de la proteína fue construida a partir de una comprobación manual de las cadenas laterales y de las moléculas del solvente. La estructura fue refinada y validada mediante el uso de *Coot* (Emsley *et al.*, 2010), *phenix.refine* (Afonine *et al.*, 2010) y *MolProbity* (Chen *et al.*, 2010). En la tabla 4.12 se muestran las estadísticas obtenidas de estos procesos.

	<i>EcOPRT</i> -5FOA	<i>EcOPRT</i> -5AOA
PDB code	-	-
Data collection		
Wavelength (Å)	0.9801	0.9801
Resolution range (Å)	42.57-1.435 (1.49-1.43)	42.7-1.62 (1.68-1.62)
Space group	<i>P</i> 2 ₁ 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁ 2 ₁
a, b, c (Å)	53.68, 69.88, 104.16	53.74, 70.31, 104.51
β (°), $\alpha = \gamma = 90^\circ$	90, 90, 90	90, 90, 90
Total reflections	252,922 (24,930)	240,695 (23,276)
Unique reflections	71,394 (7,047)	50,723 (4,936)
Multiplicity	3.5 (3.5)	4.7 (4.7)
Completeness (%)	99.00 (98.00)	99.00 (98.00)
Mean I/ σ (I)	18.46 (3.66)	20.38 (3.35)
Wilson B-factor	14.98	18.53
R-merge	0.036 (0.316)	0.044 (0.457)
R-meas ^a	0.043	0.050
CC1/2 ^b	0.999 (0.923)	0.999 (0.89)
CC ^c	1 (0.98)	1 (0.97)
Refinement		
R-work	0.1301 (0.1397)	0.1773 (0.2289)
R-free	0.1675 (0.2048)	0.2077 (0.2508)
# non-H atoms	3,859	3,796
# Protein	3,391	3,399
# Ligand atoms	66	60
# Water	402	694
Protein residues	419	421
RMS (bond) (Å)	0.022	0.007
RMS(angles)(°)	1.88	1.10
Ramachandran analysis		
Favored/Allowed/Outlier (%)	99.0/0.93/0.0	99.0/2.71/0.0
Clashscore	8.40	4.04
Average B-factor (Å ²)	22.54	25.20
Protein	20.86	24.05
Ligands	29.16	39.92
Solvent	35.58	34.21

Tabla 4.12. Datos estadísticos de la solución de las estructuras de *EcOPRT* acomplejada con 5-FOA y 5-AOA.

4.10.11. Estructura cristalográfica de *EcOPRT* en complejo con los análogos 5-FOA y 5-AOA

En línea con las estructuras de otras fosforribosiltransferasas (PRTasa) tipo I, la estructura de *EcOPRT* consiste en un largo β -hairpin N-terminal llamado “tapa” (“hood”;

Resultados

hebras β 1-4) anclado por una hélice α , un plegamiento Rossmann central donde destaca el largo lazo catalítico, y dos hélices α C-terminales. En total, la topología de *EcOPRT* incluye 8 hélices α y 9 hebras β . *EcOPRT* es un dímero en solución. La “tapa”, junto con la hélice α 2, ayuda a mantener la posición del sustrato natural orotato (OA, ácido orótico) (figura 4.54).

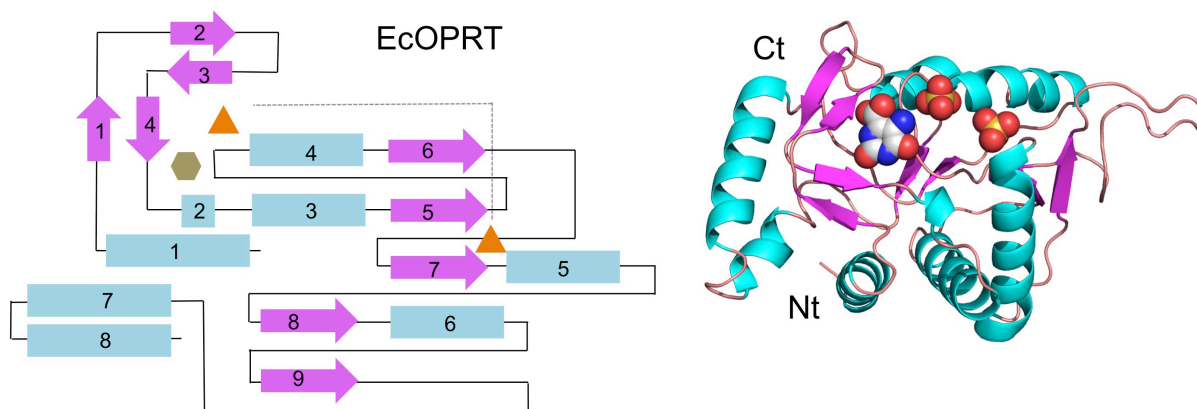


Figura 4.54. Esquema de la estructura secundaria de *EcOPRT*. A la izquierda el esquema con las hélices α en azul y las hebras β en violeta. Como triángulos naranjas aparecen representados los átomos de sulfato y en hexágono verde, el ángulo y la posición de los ligandos representados en modelo de bolas y coloreados según el átomo (en la figura aparece representada la molécula de 5-AOA). A la derecha representación la estructura tridimensional del monómero de *EcOPRT* acomplejada con 5-AOA.

Para entender las implicaciones funcionales y estructurales de los análogos de sustrato 5-amino-orotato (5-AOA) y 5-fluoro-orotato (5-FOA), hemos resuelto las estructuras cristalográficas de los dos complejos de *EcOPRT* con 5-AOA y con 5-FOA, a 1,62 Å y 1,43 Å de resolución, respectivamente. En ambos complejos, una molécula del análogo de sustrato correspondiente se localizó por centro activo de *EcOPRT*. La estructura en complejo con los análogos 5-AOA y 5-FOA no mostraron cambios conformacionales en comparación con la estructura resuelta previamente en ausencia de sustratos (pdb 1oro) (r.m.s.d. de 0,87 Å y 0,79 Å para la superposición apo/5-AOA y apo/5-FOA, respectivamente, utilizando los dímeros completos). Ambas moléculas de análogos estructurales se posicionan entre la tapa y la hélice α 2 (figura 4.55).

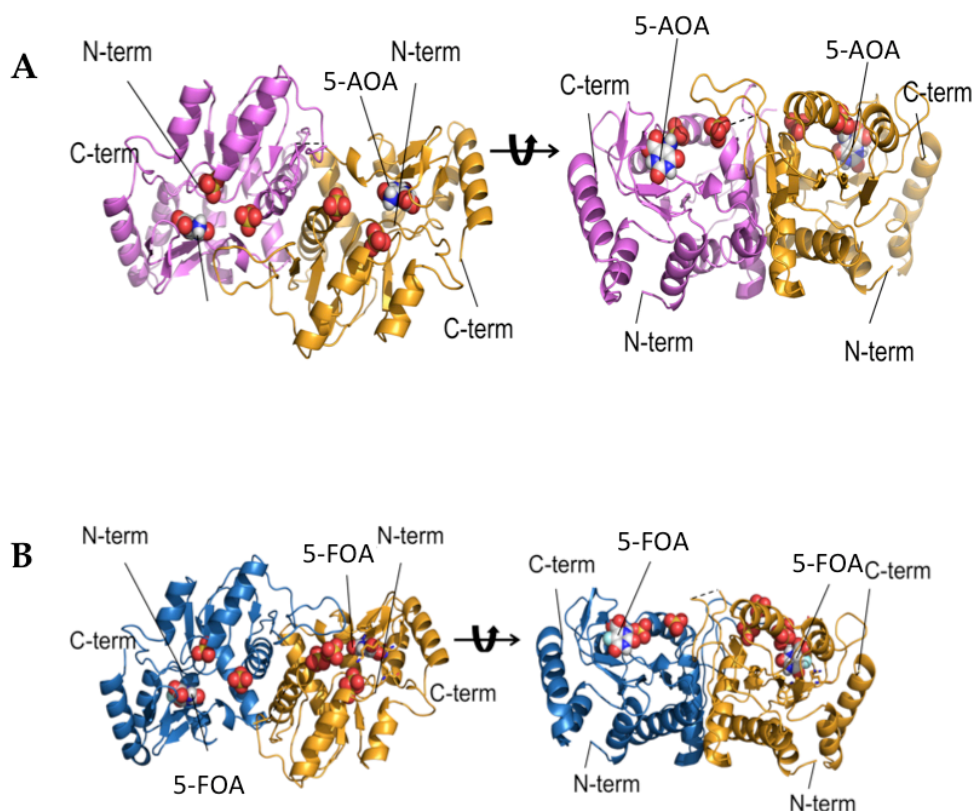


Figura 4.55. Estructura tridimensional de los complejos de EcOPRT con 5-AOA y 5-FOA.

A) Representación en dos orientaciones del complejo de EcOPRT con 5-AOA. Los ligandos se hallan representado en modelo de bolas. **B)** Representación en dos orientaciones del complejo de EcOPRT con 5-FOA. En ambas figuras la orientación relativa de ambas orientaciones es de 90°.

Dada la alta resolución a la que difractaron los cristales de ambos complejos de EcOPRT, la densidad electrónica permitió definir sin ambigüedades la posición de cada una de las moléculas (véase figura 4.56). Los residuos del centro activo implicados en la unión son el O del grupo carbonílico de la cadena principal de Phe35, el N de la cadena principal de Phe34 y Lys26, y la cadena lateral de Arg156.

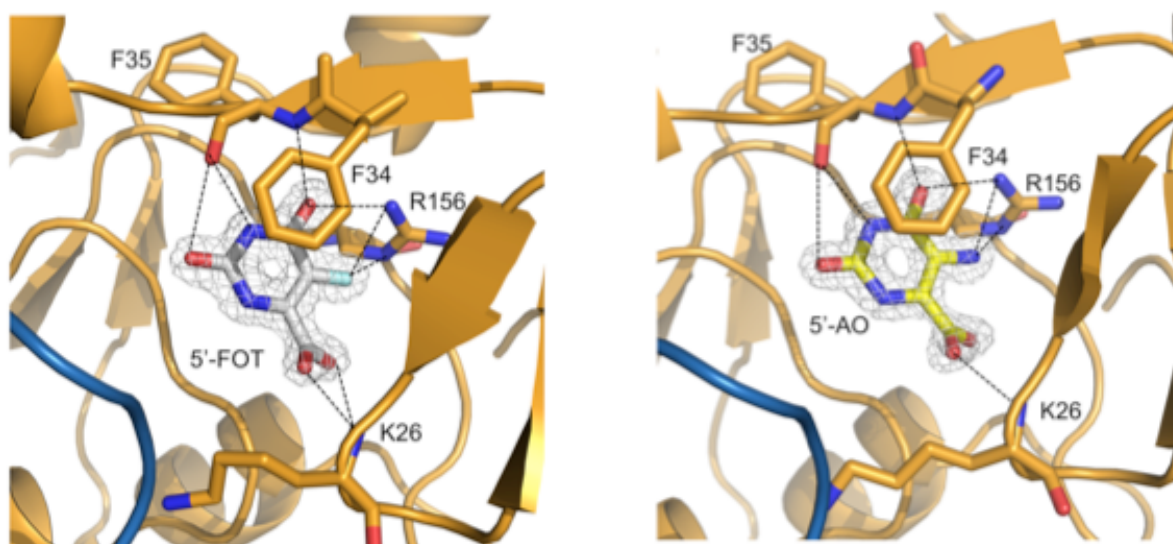


Figura 4.56. Densidad electroónica de las moléculas de 5'FOT (5-FOA) y 5'-AO (5-AOA) en el centro activo de *EcOPRT*. En ambas figuras el mapa de densidad electrónica (2Fo-Fc) se halla contorneada a 1σ .

4.10.11.1. Presencia de sulfatos en el centro activo de *EcOPRT*

En los complejos de *EcOPRT* con 5-FOA y 5-AOA se puede observar la presencia de dos moléculas de iones sulfato por centro activo, que ocupan siempre dos posiciones conservadas. Estas posiciones corresponden a los sitios de unión de dos de los grupos fosforilo del sustrato PRPP, por lo que los iones sulfato cumplirían una función inhibitoria de la enzima al competir con el PRPP. En la figura 4.57 se puede observar como uno de las moléculas de sulfatos se halla en la posición del *S-loop* interaccionando con los residuos Thr128-Ala132 y con los residuos de Lys26 y Lys73 del mismo monómero. El residuo Lys73 participa en la unión de ambos sulfatos a través de ambas cadenas, lateral y principal, respectivamente. Adicionalmente, la segunda molécula de sulfato se halla interaccionando con residuos procedentes de ambas subunidades, Arg100 y Tyr72 del mismo monómero y las cadenas laterales de los residuos Arg100 y Lys103 del segundo monómero.

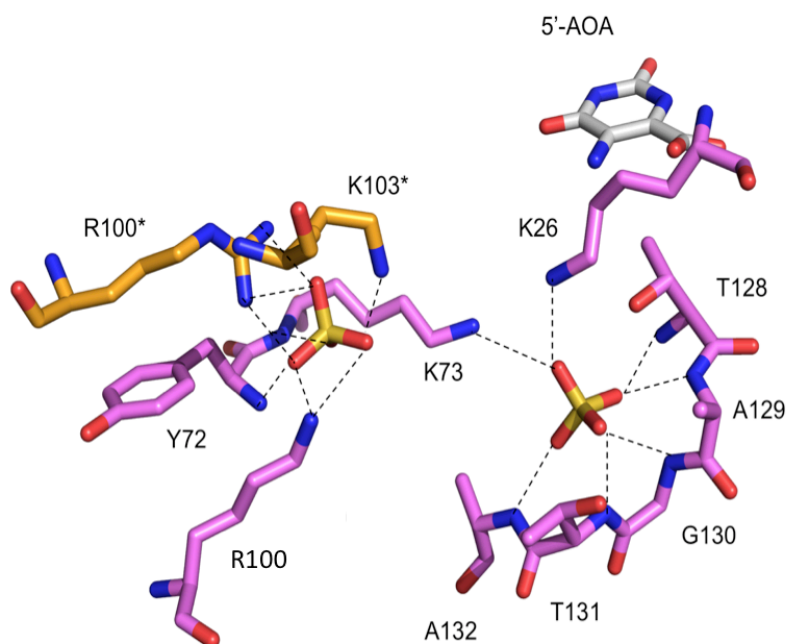


Figura 4.57. Posición de dos moléculas de sulfato en el centro activo de *EcOPRT*. Este tipo de moléculas actuarían como inhibidores de la actividad enzimática.

4.10.12. Resolución estructural y refinamiento de *MtOPRT* en su forma apo

Todos los datos de la estructura apo fueron integrados con XDS (Kabsch, 2010) y escalados gracias a *Aimless* (Evans y Murshudov, 2013). La estructura se resolvió mediante la técnica de MR, gracias a la estructura apo (pdb 1oro) (Henriksen *et al.*, 1996) con la ayuda de PHASER (McCoy *et al.*, 2007). La estructura fue construida a partir de una comprobación manual de las cadenas laterales y de las moléculas del solvente. La estructura fue refinada y validada mediante el uso de *Coot* (Emsley *et al.*, 2010), *phenix.refine* (Afonine *et al.*, 2012) y *MolProbity* (Chen *et al.*, 2010). En la tabla 4.13 se muestran las estadísticas obtenidas de estos procesos.

<i>MtOPRT</i>	
PDB code	-
Data collection	
Wavelength (Å)	1.031
Resolution range (Å)	43.53-1.54 (1.595-1.54)
Space group	<i>P</i> 1 2 ₁ 1
a, b, c (Å)	52.5, 66.9, 53.9
β (°), $\alpha = \gamma = 90^\circ$	90, 110, 90
Total reflections	132,607 (10,052)
Unique reflections	49,426 (4,682)
Multiplicity	2.7 (2.1)
Completeness (%)	0.95 (0.90)
Mean I/ σ (I)	14.7 (3.75)
Wilson B-factor	11.76
R-merge	0.071 (0.238)
R-meas ^a	0.085 (0.347)
CC1/2 ^b	0.982 (0.818)
CC ^{*c}	0.996 (0.949)
Refinement	
R-work	0.1570 (0.1968)
R-free	0.1907 (0.2230)
# non-H atoms	3,283
# Protein	2,879
# Ligand atoms	33
# Water	371
Protein residues	372
RMS (bond) (Å)	0.014
RMS(angles)(°)	1.42
Ramachandran analysis	
Favored/Allowed/Outlier (%)	98.0/1.9/0.53
Clashscore	10.79
Average B-factor (Å ²)	21.05
Protein	19.41
Ligands	42.59
Solvent	31.86

Tabla 4.13. Datos estadísticos de la solución de la estructura de *MtOPRT* en su forma apo.

4.10.13. Estructura tridimensional de *MtOPRT* en su forma apo

El plan estructural de *MtOPRT* corresponde a las PRTas de tipo I, como las enzimas de *E. coli* y *S. typhimurium*, a pesar de la baja similaridad de secuencia entre ellas. En particular, la secuencia de *MtOPRT* es más corta que las secuencias equivalentes (176 aminoácidos frente a 211 en *EcOPRT*), resultado de la ausencia en *MtOPRT* de los últimos 35 aminoácidos presentes en *EcOPRT*. El motivo “tapa” o “hood”, formado por las primeras 4 hebras β , se encuentra muy bien definido en *MtOPRT*, así como el plegamiento Rossmann central; en cambio, de las dos últimas hélices α C-terminales de *EcOPRT* únicamente queda una vuelta de hélice en *MtOPRT*. En total, los elementos de estructura secundaria de *MtOPRT* son 9 hebras β y 7 hélices α (figura 4.58 y 4.59). Como *EcOPRT*, *MtOPRT* es un dímero en solución.

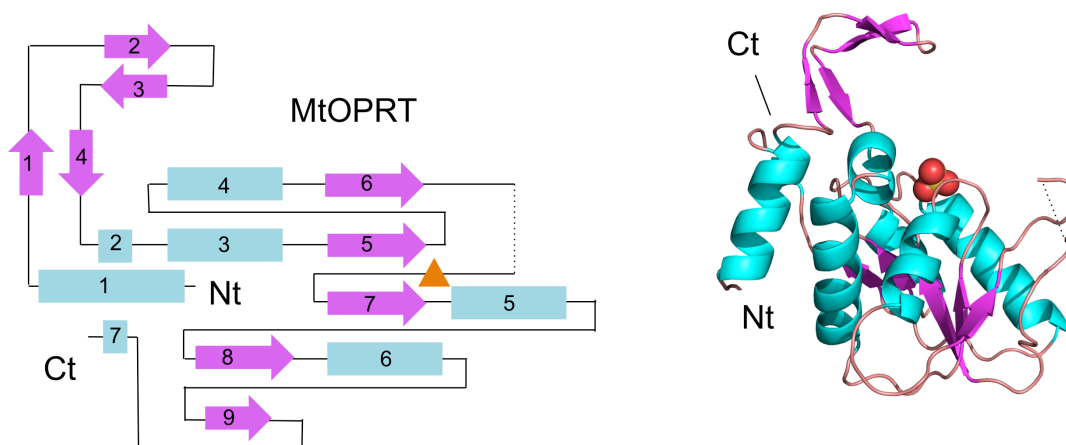


Figura 4.58. Esquema de la estructura secundaria de *MtOPRT*. A la izquierda el esquema donde se representan las hélices α en azul y las hebras β en violeta. Como triángulos naranjas aparecen representados los átomos de sulfato. **B)** Representación la estructura tridimensional del monómero de *MtOPRT*. Las dos moléculas de sulfatos por centro activo se representan en modelo de bolas y coloreados según el átomo.

El lazo catalítico que durante la catálisis cubre la cavidad del sitio activo del monómero vecino se encuentra desordenado en la estructura de *MtOPRT* (figura 4.58); dada la alta resolución de los datos cristalográficos, y la situación análoga que existe con uno o los dos lazos catalíticos en la estructura de *EcOPRT*, inferimos que el

Resultados

desorden observado responde a la flexibilidad intrínseca de este lazo y que dicha flexibilidad está involucrada funcionalmente en ensamblar los sitios activos compuestos haciendo así posible la catálisis.

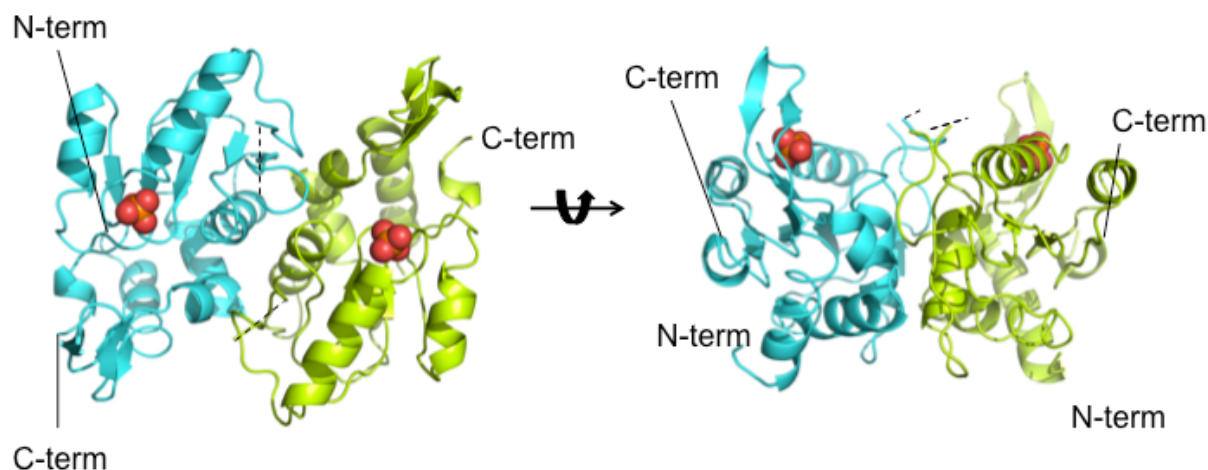


Figura 4.59. Estructura tridimensional de *MtOPRT*. Orientaciones relacionadas por 90° del dímero *MtOPRT* representando ambos monómeros con cian y verde. La molécula de sulfato se halla representada en bolas.

En el centro activo de *MtOPRT* se encuentran iones sulfato procedentes de la condición de cristalización, que se sitúan en las posiciones que deberían ocupar los grupos fosforilo- β del PRPP; concretamente, los iones sulfato se localizan junto al motivo *S-loop* entre las hebras β 7 y la hélice α 5. La interacción del sulfato con *MtOPRT* implica interacciones tanto con grupos amida de la cadena principal (Thr124, Thr125, Gly126) como enlaces de hidrógeno con grupos hidroxilo cercanos (Gly126, Ser128) (figura 4.60). Dada la fuerte interacción que existe con ellos, los iones sulfato actuarían como inhibidores competitivos del PRPP, desplazando los grupos fosfato del sustrato.

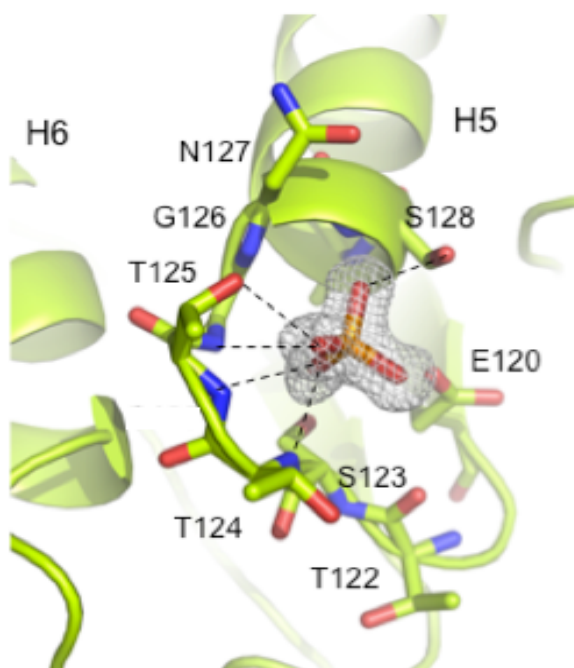


Figura 4.60. Detalle de la molécula de sulfato en el S-loop en *MtOPRT*. Este tipo de moléculas encontradas en el centro activo actuarían como inhibidores de la reacción enzimática.

4.10.14. Superficie de dimerización de *MtOPRT*

La mayor parte de la superficie de dimerización de *MtOPRT* está compuesta por las hélices α 2, 3 y 4 y la hebra β 6 contigua. La superficie total encerrada es de 1.130 Å² (11,4% de toda la superficie expuesta en *MtOPRT*), incluye 32 residuos, y está formada por 9 enlaces de hidrógeno y 7 puentes salinos. Es una superficie de interacción extensa que asegura la estabilidad del dímero, además de cumplir un papel funcional relevante. Los dos sitios activos de la OPRT no actúan de forma independiente sino que, al tratarse de sitios activos muy expuestos al solvente y muy abiertos, como consecuencia del lazo catalítico desordenado, cuando los sustratos están presentes y ha de tener lugar la catálisis, el lazo catalítico de la subunidad adyacente se extiende sobre el sitio activo de la otra subunidad, ordenándose en una estructura de tipo hebra β , donando residuos críticos para la reacción. Por tanto, el dímero de *MtOPRT* es un oligómero obligado con centros activos compuestos por residuos de ambas subunidades.

Discusión

5. Discusión

5.1. Mecanismo de transpersulfuración en el sistema CSD de *E. coli*

La secuencia motivo de las cisteína desulfurasas de tipo II, a la que pertenecen tanto CsdA como SufS, engloba la región 356-QHCAQ-360 alrededor de la cisteína catalítica, Cys358. Esta secuencia motivo está conservada y es diagnóstica de este clase de cisteína desulfurasas. En efecto, la clase I, a la que pertenece IscS, posee una secuencia consenso diferente. A nivel estructural, la secuencia motivo de las cisteína desulfurasas de clase I adquiere una estructura flexible y móvil, coherente con su función; en cambio, la estructura de la secuencia motivo de la clase II adopta una estructura definida que no sufre cambios relevantes en su configuración durante las reacciones de persulfuración o transpersulfuración. Sin embargo, más allá de la conservación de la secuencia-estructura, antes de este trabajo no se sabía si existían papeles catalíticos que justificasen la conservación de estos motivos.

La reacción de transferencia del grupo persulfuro implica un ataque nucleofílico (Ayala-Castro *et al.*, 2008) por parte del anión tiolato de la cadena lateral de una cisteína (Cys61 de CsdE) sobre el azufre persulfuro (CsdA Cys358 $S \gamma -S \delta$), resultando en la formación de un nuevo persulfuro en CsdE Cys61 y liberándose el tiol de CsdA Cys358.

El trabajo de esta tesis ha contribuido a establecer las bases estructurales y funcionales de la transferencia de azufre a través de interfaces proteína-proteína, utilizando el complejo CsdA-CsdE como modelo. La estructura cristalográfica de (CsdA_{per}-CsdE_{per})₂ a 2,49 Å de resolución (Figura 5.1a) demuestra por primera vez cuál es el rol del motivo conservado de las cisteína desulfurasas de clase II en el mecanismo de la transpersulfuración (Fig. 5.1b). En la estructura tanto CsdA como CsdE están persulfuradas, situación que se daría al final de la transpersulfuración, asumiendo que CsdA Cys358 pueda re-persulfurarse antes de que CsdE se disocie del complejo. En concreto, los grupos amida del esqueleto polipeptídico entre Gln356 y Gln360 están alineados y crean una concentración local de carga positiva a lo largo del trayecto entre el azufre persulfuro de CsdA Cys358 y el azufre persulfuro de CsdE Cys61. Esta concentración local de carga actuaría como un hueco catiónico, estabilizando la formación de intermedios cargados negativamente en esa zona. Dado que la Cys61 de CsdE tiene su cadena lateral desprotonada y que el ataque nucleofílico genera una carga negativa naciente en el azufre persulfuro de CsdA Cys358, la

Discusión

estabilización aportada por los residuos que forman parte del motivo cisteína desulfurasa resultaría clave para facilitar la catálisis.

En la estructura de (CsdAper-CsdEper)₂ la distancia que separa los átomos de azufre implicados en la reacción, CsdA Cys358 S γ -S δ y CsdE Cys61 S S γ es de 4,5 Å. Esta distancia es compatible con un ataque nucleofílico. Es interesante considerar que en la estructura del complejo sin modificar (Kim y Park, 2013) la posición de la cadena lateral de CsdE Cys61 no se encontraba modelada y, por tanto, no permite establecer distancias de enlace ni realizar predicciones sobre la geometría de reacción. La estructura de (CsdAper-CsdEper)₂, en cambio, posee las cadenas laterales de las dos cisteínas catalíticas bien definidas en densidad electrónica, lo cual posibilita realizar inferencias sobre el mecanismo. Con pequeños ajustes, los enlaces CsdA Cys358 S γ -S δ y CsdE Cys61 C β -S γ podrían quedar alineados y a distancia de enlace. Además de los residuos del motivo de cisteína, las cadenas laterales de la His120 y Lys250* (* significa que el residuo proviene de la subunidad opuesta), además del grupo amida del enlace peptídico de la propia Cys358 (Figura 5.1). En cuanto a CsdE, además de la Cys61, el grupo carboxilato de la cadena lateral de Glu62 interacciona con CsdA durante la formación del complejo.

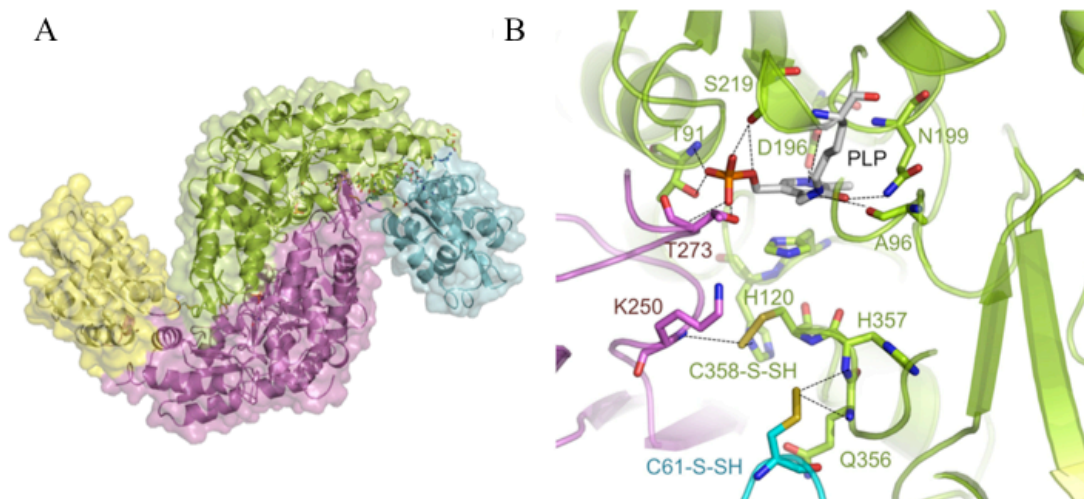


Figura 5.1. Representación de la estructura del complejo persulfurado (CsdAper-CsdEper)₂. **a** Representación del complejo persulfurado. CsdA se muestra en verde y violeta, mientras que CsdE aparece en amarillo y azul. **b** Detalle estructural de los residuos implicados en el proceso de transpersulfuración, indicando la molécula de PLP, la Cys358 (CsdA) y la Cys61 (CsdE) en colores CPK. La interacción estabiliza el átomo de azufre del persulfuro (las interacciones se marcan en línea discontinua) (Fernández *et al.*, 2016).

La conservación general de la secuencia y estructura entre las cisteína desulfurasas de clase II indica que el mecanismo de transpersulfuración de SufS, que es homóloga de CsdA (identidad de secuencia es 43%), esté también conservado. La homología estructural entre SufS y CsdA es bastante clara en superposiciones de las dos estructuras (Figura 5.2). Además, SufE, la proteína aceptora de azufre de SufS, es homóloga de CsdE. Esta doble homología sugiere que podrían existir fenómenos de *crosstalk* entre los dos sistemas, SUF y CSD, es decir, que CsdA pudiera transferir azufre a SufE y SufS a CsdE, al menos bajo determinadas circunstancias. Las diferencias estructurales entre CsdA y SufS son pequeñas y están limitadas a elementos de la zona N-terminal próximos al lugar de unión de CsdE/SufE.

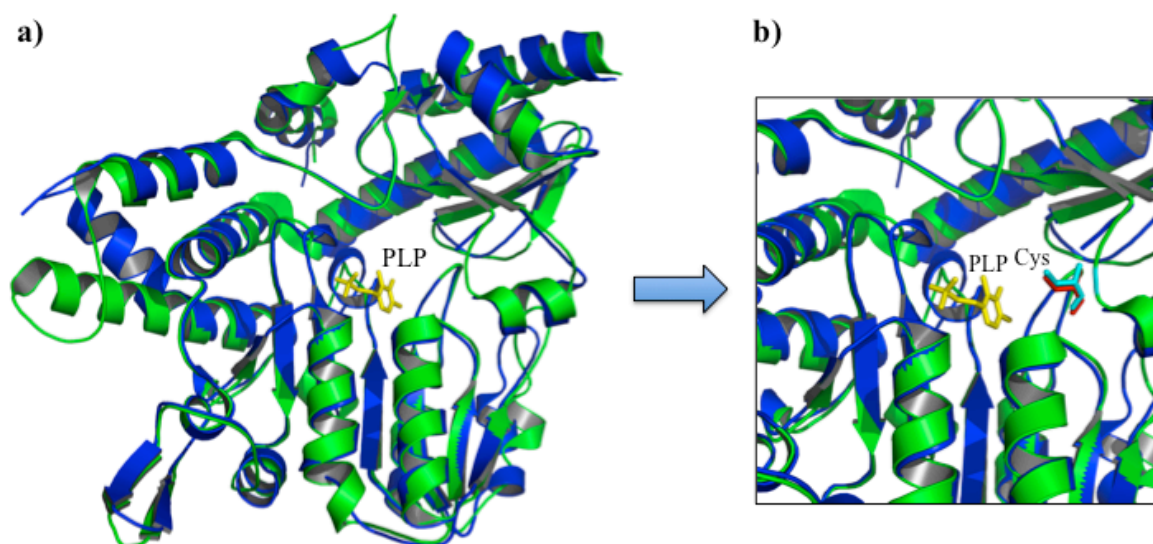


Figura 5.2. Alineamiento estructural de las dos cisteína desulfurasas de *E. coli*, CsdA y SufS. a) La estructura en color verde corresponde a SufS, mientras que la de color azul marino es CsdA. Se observa que el alineamiento es prácticamente perfecto. b) Detalle del centro activo donde se puede apreciar el cofactor PLP (molécula amarilla) y las cisteínas capaces de realizar los procesos de persulfuración y transpersulfuración. La moléculas en rojo corresponde al residuo Cys358 de CsdA y la de color cian a la cisteína Cys364 de SufS. Los códigos pdb fueron 1c0n para SufS y 4lw2 para CsdA.

En la figura 5.2 además se puede observar cómo los elementos fundamentales para la persulfuración y la transpersulfuración se conservan en las dos enzimas, al alinearse estructuralmente las moléculas de PLP y las cisteínas catalíticas de cada una de las enzimas.

5.2. La interacción de CsdA y CsdE es una unión dinámica

La interacción entre CsdA y CsdE es fuerte y tiene una K_d de 120 nM (Fernández et al., 2016). Sin embargo, la estructura de $(CsdA_{per}-CsdE_{per})_2$ revela que la Cys61 y residuos inmediatamente adyacentes (especialmente, Glu62) de CsdE son los que más íntimamente se unen al sitio activo de CsdA, mientras que el resto de la estructura puede adoptar muchas posiciones equivalentes (Figura 5.3). La interacción entre CsdE Cys61 y el sitio activo de CsdA define un punto de pivote alrededor del cual CsdE puede oscilar en un ángulo de 8° . Este dinamismo del complejo entre CsdA y CsdE depende en parte de la naturaleza de la superficie de interacción, que está caracterizada por un gran número de interacciones mediadas por carga (puentes salinos y enlaces

fuertes de hidrógeno). A diferencia de complejos con superficies de interacción hidrofóbicas, la naturaleza hidrofílica/cargada de la interacción CsdA-CsdE hace que ambas proteínas puedan asociarse-disociarse rápidamente.

Para confirmar la importancia de los grupos cargados en la interacción CsdA-CsdE y analizar qué interacciones en particular contribuían más a la estabilidad del complejo, realizamos mutantes simples, dobles y triples de grupos cargados de la superficie de CsdE, identificados a través del análisis de la estructura de (CsdAper-CsdEper)₂. Las interacciones entre CsdA y los mutantes de CsdE fueron analizadas mediante SPR. A pesar de que los mutantes sencillos R86E y R129E de CsdE no afectaron excesivamente la asociación, los dobles (CsdE Y34E/R35E y CsdE K42E/R86E) y triples mutantes (CsdE Y34E/R35E/R129E y CsdE K42E/R86E/R129E) esencialmente abolieron la interacción (K_d = 29.940-52.000 nM para los dobles y no determinable para los triples). En contraste, un mutante doble control en la cara opuesta a la de interacción (R21E/N22E) tenía una K_d (177 nM) prácticamente idéntica a la proteína nativa (125 nM, estimada también por SPR).

El trabajo de esta tesis ha permitido identificar los residuos clave para la interacción, tanto a nivel estructural (interfaces) como a nivel de residuos individuales, y establecer que son las interacciones mediadas por carga/electrostáticas las más importantes. Es interesante que en la estructura del complejo apo CsdA-CsdE los residuos identificados como importantes para la interacción no se superponen perfectamente a los indentificados en el complejo persulfurado, y las interacciones electrostáticas no se consideraron tan importantes. Es posible que cambios en la conformación de CsdE entre ambos complejos pueda explicar, al menos en parte, las diferencias observadas en las interfaces de interacción CsdA-CsdE.

5.3. CsdA es capaz de interaccionar con componentes de los sistemas CSD y Suf

Por otro lado, mediante los experimentos de MALDI-TOF, y la persulfuración del complejo (CsdA-CsdE)₂, se ha podido observar que el sistema CSD es capaz de transferir átomos de azufre al sistema Suf, en concreto a la proteína SufE, tanto desde el complejo (CsdAper-CsdEper)₂ como desde (CsdAper)₂. Cuando se produce este fenómeno parecen generarse tres especies de SufE, cargada con 1, 2 ó 3 átomos de azufre, según los datos obtenidos, en el caso de la persulfuración vía el complejo. Sin embargo, vía (CsdAper)₂ se produce solo una especie de SufE, cargada con un solo átomo de S. Por esta razón el proceso conocido como *crosstalk* se puede considerar un indicativo de la interacción dinámica entre los componentes del sistema CSD y Suf de *E. coli*, pues bajo determinadas condiciones ambientales esa dinámica puede ser esencial para el metabolismo celular. Estos hechos además son reforzados por el gran parecido

estructural de los componentes de los dos sistemas, CsdA/SufS, así como el par CsdE/SufE (figura 1.8 y 5.2), por lo que el *crosstalk* es explicable estructuralmente. Sin embargo, hasta esta tesis no se ha podido detectar la transferencia directa de azufre desde CsdA a SufE, por lo que marca la evidencia de que CsdA puede actuar para los dos sistemas en *E. coli*, donde parece clara la participación de CsdA en el sistema Suf, al restaurar la sobreexpresión del gen *csdA* la delección del gen *sufS* (Trotter *et al.*, 2009).

En cuanto a la comparación de las proteínas aceptoras de azufre, CsdE y SufE, existen cambios en la posición de sus cisteínas catalíticas, mostrándose la de SufE (Cys51) más escondida en el bolsillo hidrofóbico del centro activo, mientras que el residuo de CsdE (Cys61) se encuentra más expuesto. Este hecho puede explicar porque una de las proteínas sigue una ruta hacia modificación de ARNt entre otros (sistema CSD) y la otra hacia la formación de *clusters* [Fe-S] (sistema Suf). Además aunque sea posible por parte de CsdA seguir ambas rutas (CSD o Suf), la función básica de CsdA parece estar dirigida hacia el sistema CSD, ya que el complejo (CsdA-CsdE)₂ es más estable termodinámicamente, al haber sido cristalizado y detectado mediante filtraciones en gel, hecho no observado en el caso en la interacción CsdA-SufE.

En cuanto a la estructura del complejo (CsdAper-CsdEper)₂, ésta aporta información de la forma de interaccionar entre las proteínas implicadas. En la estructura se observa que las hélices $\alpha 1$ y 2 de CsdE se encuentran casi perpendiculares a la hélice $\alpha 2$ de CsdAper. Esto ocurre gracias al reordenamiento de las hélices $\alpha 7$ y 8 de CsdE en el complejo. El *loop* que conecta estas dos hélices α de CsdE sufre un cambio conformacional ausente en el caso de la proteína libre, el cual ayuda a mantener el espacio y la inclinación entre estas dos hélices α . Este cambio conformacional permite el ensamblaje del complejo, aunque las cisteínas catalíticas (Cys358 en CsdA y Cys61 en CsdE) no se encuentran suficientemente próximas en el caso del complejo apo (9,1 Å) (Kim y Park, 2013), para permitir la cesión del átomo de S. En el caso de nuestra estructura persulfurada (CsdAper-CsdEper)₂, la distancia entre las proteínas es menor, encontrándose el persulfuro a una distancia de aproximadamente 4,5 Å del otro residuo de cisteína. A esta distancia la transferencia del átomo de azufre es altamente probable (figura 5.1).

Centrándonos en CsdE, ésta posee una mayor flexibilidad en su estructura ya que mientras la Cys61 se encuentra fijada en el complejo (CsdA-CsdE), el resto de CsdE es capaz de balancearse sobre la CsdA. Cuando CsdE se encuentra persulfurada (CsdEper), al interaccionar con CsdA, ésta sufre un cambio donde se da un balanceo de hasta 8° sobre un eje definido por CsdAper, en concreto sobre la hélice α en los residuos

338-349. Este grado de plasticidad estructural tanto en CsdE libre como en (CsdA-CsdE)₂ facilita el transporte de azufre dentro del sistema CSD (figura 5.3).

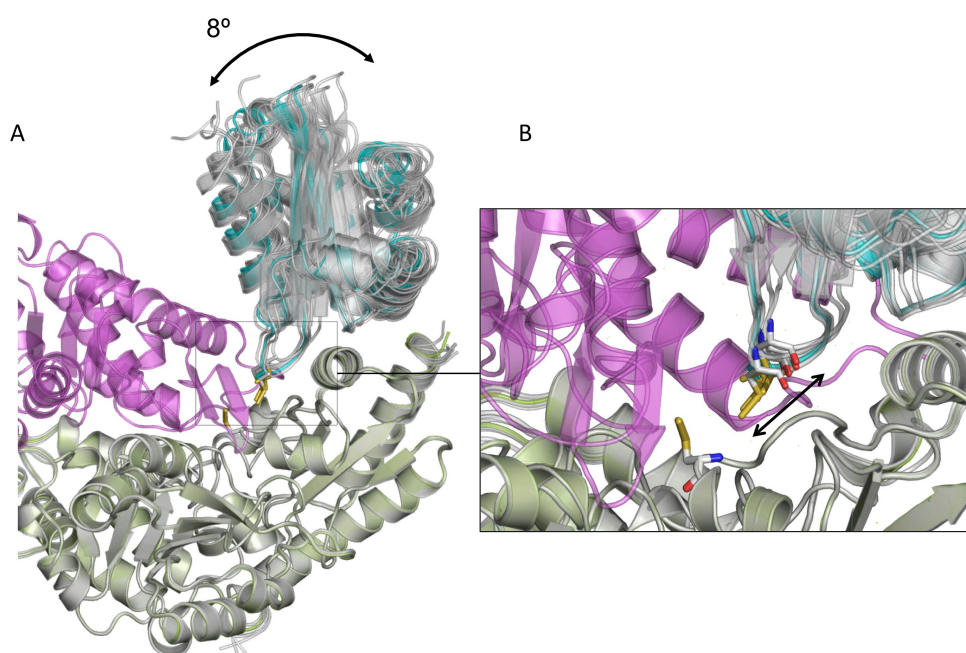


Figura 5.3. Variación estructural en la interfaz en el complejo CsdA-CsdE. **a** Superposición de las estructuras de los tetrámeros de (CsdAper-CsdEper)₂. CsdA es representada en violeta y verde y CsdE en cian. Se observa el cambio en 8° de CsdE sobre la hélice- α de CsdA. **b** Acercamiento de la interfaz CsdA-CsdE, mostrándose la interacción directamente relacionada con el movimiento de CsdE alrededor del eje definido por CsdAper, con los residuos de la hélice- α 338-349. Este hecho facilita el acercamiento de la Cys358 de CsdA hacia la Cys61 de CsdE (Fernández *et al.*, 2016).

5.4. Interacción transitoria entre CsdE y CsdA/TcdA

Fue posible observar interacciones descritas previamente entre CsdA-TcdA y TcdA-CsdE (Trotter *et al.*, 2009) por técnicas de RMN 1D y 2D HSQC, capaces de detectar interacciones transitorias con una alta sensibilidad.

En el caso de CsdA-TcdA se midieron espectros 1D ¹H-¹³C y ¹H-¹⁵N HSQC empleando TcdA doblemente marcada (¹³C y ¹⁵N) y un exceso 10 veces superior de CsdA sin marcar. Se detectó, en el caso de ¹H-¹⁵N HSQC una disminución de la intensidad en la resonancia de TcdA en presencia de CsdA, lo que sugiere la existencia

de una interacción débil, pero detectable entre CsdA y TcdA, consistente con lo observado en publicaciones anteriores (Trotter *et al.*, 2009), donde se observa la transferencia de átomos de azufre desde CsdA a TcdA. Este hecho implica necesariamente un contacto. En cuanto al experimento ^1H - ^{13}C HSQC, también se detectaron diferencias significativas aunque no tan evidentes como en el caso anterior.

Por otro lado, la detección de la interacción TcdA-CsdE fue más fácil de observar, debido a la resolución de la estructura de CsdE mediante RMN (bmr5630, pdb 1ni7, Liu *et al.*, 2005). Fue posible caracterizar la zona de CsdE capaz de interaccionar con TcdA, llevando a cabo experimentos de tipo ^1H - ^{15}N 2D HSQC (NOESY) con CsdE doblemente marcada (^{13}C y ^{15}N), en presencia o no de un exceso molar cuatro veces superior de TcdA sin marcar. Se detectaron varios picos de CsdE que sufren perturbaciones químicas en el caso de la mezcla de las proteínas, hecho no observado en el caso del análisis de CsdE sola. Esto demuestra la existencia de una interacción física transitoria entre estas dos proteínas. Además lo que se observa es un rápido proceso de intercambio, lo que explica la ausencia de la caída de la señal.

Este experimento además permitió mapear la zona de CsdE que interacciona con TcdA, siendo la zona de mayor perturbación la correspondiente a los residuos 70-81 de CsdE, muy cercana a la Cys61, residuo catalítico de CsdE. En adición, gracias a las estructuras cristalinas de CsdE y (CsdA-CsdE)₂ se sabe que esta zona es muy dinámica, permitiendo tanto la interacción con CsdA como con TcdA.

Para obtener más información de la interacción TcdA-CsdE se llevaron a cabo experimentos con el entrecruzador específico de grupos tiol, BMOE. Fue posible, gracias al análisis por PAGE-SDS observar el entrecruzamiento entre un monómero de TcdA y otro de CsdE, no observándose la formación de especies entrelazadas en el caso del análisis de TcdA sola o CsdE sola. Este hecho implica por parte de CsdE necesariamente la Cys61 (único residuos de cisteína en la secuencia de CsdE). En cuanto a la cisteína de TcdA capaz de recibir el átomo de S por parte de CsdE, gracias a los experimentos de entrecruzamiento, la tripsinización y los análisis por MALDI-TOF, parece que la recepción del átomo solo es posible o por parte de la Cys203 o la Cys220 de TcdA (figura 5.4). Sin embargo, la estructura cristalina de TcdA muestra que la Cys203 se encuentra enterrada en la estructura, mientras que la Cys220 está expuesta. Este hecho convierte a esta última Cys220 en la candidata a recibir los átomos de azufre.

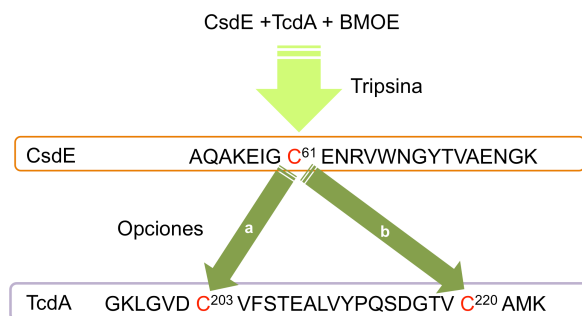


Figura 5.4. Esquema de las posibles interacciones a través de la molécula de BMOE entre TcdA y CsdE. Se muestra como la interacción y la hipotética transferencia de azufre se daría desde la Cys61 de *E. coli*, hacia la Cys203 o la Cys220 de TcdA.

Este complejo formado entre TcdA-CsdE gracias al BMOE, pudo ser analizado mediante SAXS, dando como resultado la formación de una envoltura que sugiere que el dímero de TcdA se encuentra en el centro del complejo, con una molécula de CsdE a cada lado en posiciones pseudosimétricas, resultado obtenido también gracias a la solución de las estructuras cristalinas de TcdA y CsdE. De esta forma los datos de dispersión obtenidos fueron procesados e incorporados juntos en base a un radio de giro (Rg) coherente con los datos obtenidos. Se obtuvieron los siguientes datos: Rg del complejo TcdA-CsdE 39 Å (puntos de Guinier 24-53; sRg 0,471-0,930; fidelidad 0,706) y Rg 28 Å del dímero de TcdA (puntos de Guinier 5-72; sRg 0,139-0,892; fidelidad 0,979). De esta forma los datos SAXS (peso molecular del complejo 90 kDa) demuestran la existencia de una especie donde se da el dímero de TcdA junto con dos moléculas de CsdE, al compararlo con los valores teóricos (peso molecular del complejo 104 kDa) (tabla 5.1).

	Masa teórica ¹ (Da)	SAXS (Guinier) ²	SAXS (Porod) ³
CsdE	16760,1	n.a.	n.a.
TcdA-CsdE	91929,5	154919,0	128132,0

Tabla 5.1. Comparación de las masas moleculares estimadas a partir de los datos de SAXS. En todos los casos las masas estimadas coinciden con los datos esperados, corroborando la formación del complejo TcdA-CsdE (entrecruzado por BMOE).

¹Las masas teóricas fueron calculadas usando la secuencia primaria menos la metionina inicial, la cual está ausente en las proteínas recombinantes TcdA y CsdE. Las masas de 2 moléculas de ATP⁺ (2 x 503,18 Da), 2 K⁺ (2 x 39,1 Da) y una de Na⁺ (23,0 Da) fueron añadidas al dímero de TcdA, y complejo entrecruzado TcdA-CsdE. Para el complejo de entrecruzamiento de TcdA-CsdE, la masa neta añadida para el entrecruzador (BMOE, 220,05 Da) fue añadido una vez por cada molécula de CsdE en el complejo.

²Masa molecular estimada derivada de la aproximación de Guinier.

³Masa molecular estimada del volumen de Porod.

5.5. TcdA tiene la topología de la familia ThiF pero carece de actividad de adenilación y del dominio de unión a zinc

Tras resolver mediante cristalografía de rayos-X por primera vez la proteína TcdA de *E. coli* ha sido posible observar que ésta realmente se puede clasificar, estructuralmente, dentro de la familia de tipo E1 y dentro de ésta en la familia ThiF,. Existen semejanzas estructurales entre los 180 aminoácidos que componen la mitad N-terminal de TcdA y el dominio catalítico de las proteínas MoeB/ThiF, componentes de la familia ThiF. Además otras proteínas como MccB perteneciente a la superfamilia de tipo E1 de *E. coli*, así como la proteína humana UBA5 (superfamilia E1), comparten el mismo plegamiento con MoeB/ThiF, y por lo tanto con parte de TcdA (López-Esteva *et al.*, 2015).

La zona de unión de nucleótidos de TcdA, así como una gran parte de las hélices que componen cada subunidad del homodímero, participan activamente en la estabilización del complejo, correspondiendo a las estructuras más conservadas en TcdA y MoeB/ThiF (Lake *et al.*, 2001; Duda *et al.*, 2005). Además el *loop* P, el encargado de unir el ATP o AMP en el caso de TcdA, y los residuos de contacto con ATP (o AMP) se encuentran estrictamente conservados a través de los homólogos estructurales descritos hasta ahora.

En cuanto a la mitad C-terminal de TcdA, ésta forma una estructura única, que la diferencia claramente de proteínas como MoeB/ThiF. Es aquí donde en TcdA se encuentra la zona de unión del átomo de K^+ , mientras que en el par MoeB/ThiF el metal unido es Zn^{2+} (Lake *et al.*, 2001; Duda *et al.*, 2005). De las cuatro cisteínas empleadas en coordinar el átomo de Zn^{2+} en MoeB/ThiF, tan solo se conserva en TcdA la Cys220, cuya función puede estar involucrada en la captación de S. Ésta se encuentra en el comienzo del *loop* desordenado de TcdA en la estructura cristalina, pudiendo actuar como residuo catalítico durante la actividad enzimática de TcdA. Por su parte, la Cys234 de TcdA, presente en MoeB/ThiF, se encuentra en el *loop* desordenado, por lo que no es posible ver exactamente su posición. En cuanto a la Cys66 de TcdA, sin equivalencia en MoeB/ThiF, ésta se encuentra próxima al centro activo, por lo que también se puede hipotetizar sobre su posible actuación en la actividad enzimática de TcdA.

En cuanto a las diferencias entre la mitad C-terminal de TcdA y MoeB/ThiF, los motivos superficiales que permiten la interacción de MoeB con MoaD y ThiF con ThiS, se encuentran bloqueados en TcdA por el segmento comprendido en la secuencia primaria entre el aminoácido 170 y el 189, el cual forma parte de la unión al K^+ . Este segmento no aparece en MoeB/ThiF.

Al realizar alineamientos estructurales de TcdA con MoeB, ThiF, MccB y UBA5, se observan diferencias en la organización de los dominios descritos entre la hélice encontrada en el extremo C-terminal de TcdA y los dominios de unión a metales. En TcdA el sitio de unión al metal K^+ se encuentra en una posición central, entre el dominio de tipo E1 y el extremo C-terminal, donde encontramos la hélice $\alpha 8$, mientras que en el resto de los casos la zona de unión al metal (Zn^{2+}) se encuentra en el extremo C-terminal (figura 5.5).

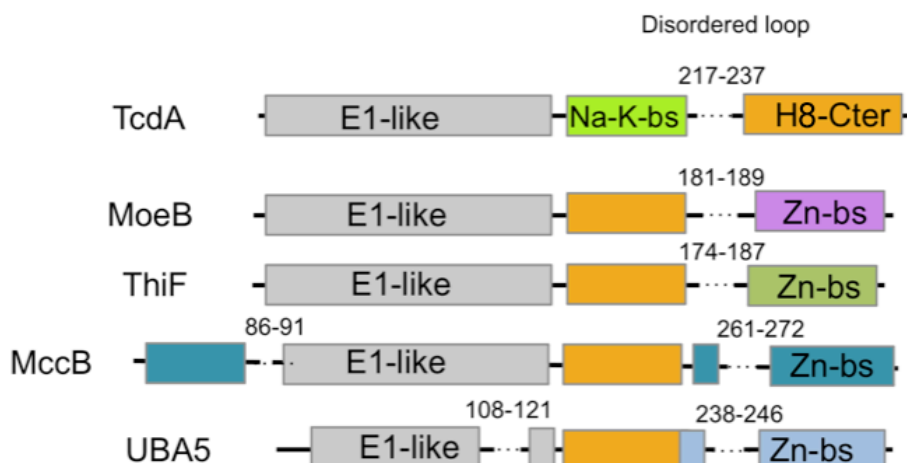


Figura 5.5. Comparación de los modelos de proteínas E1 o de tipo E1. Se representan esquemáticamente los dominios de las enzimas E1 (UBA5) o de tipo E1 (TcdA, MoeB, ThiF y MccB), con un código de colores para cada uno de los dominios. bs, sitio de unión (*binding site*); la secuencia en cada proteína correspondiente a *loops* desordenados se indican con líneas discontinuas, además de indicarse los residuos que comprenden. La hélice α 8 C-terminal de TcdA es estructuralmente equivalente a la hélice α que se encuentra junto al dominio E1 o de tipo E1 en el resto de las proteínas (López-Esteba *et al.*, 2015).

Por lo tanto, TcdA parece ser una nueva proteína, con una nueva función y que se diferencia claramente, tras la resolución de la estructura, de sus homólogos.

5.6. TcdA presenta un nuevo modo de unión a ARN^t_{Lys}(UUU), cuyo complejo ha podido ser estudiado mediante diferentes técnicas

Gracias a los ensayos realizados sobre el complejo TcdA-ARN^t_{Lys}(UUU) fue posible demostrar que esta proteína es la primera descrita dentro de su familia como capaz de interaccionar con ciertos ARNts. Los experimentos de EMSA empleando TcdA y ARN^t_{Lys}(UUU), previamente purificados, junto con ATP · Mg²⁺, permitió estudiar la interacción entre estos dos componentes (proteína y ARN). Se observó la formación de un complejo entre la proteína y el ácido nucleico, con migraciones de las bandas muy diferentes a la de los componentes por separado, siendo la proteína incapaz de penetrar en el gel debido a su pI altamente básico (8,8), en el caso de no interaccionar con el

ácido nucleico. En el caso del ARNt solo se origina una banda de peso molecular menor al caso del complejo. En el caso de la combinación del ARN con la mayor concentración de proteína ensayada (15 μ M), se da la aparición de bandas de mayor peso molecular a la que correspondería el complejo, pudiendo ser complejos de un orden superior y no el esperado con estequiometría 1:1, donde cada monómero de TcdA une una molécula de ARNt. El resto de las bandas pueden ser complejos donde un dímero de la proteína TcdA se haya unido a una sola molécula de ARN.

Los ensayos de AUC, tanto de SV como de SE demuestran con una metodología diferente la existencia del complejo entre la proteína y el ácido nucleico, y la estequiometría 1:1 (TcdA-ARNt). En el caso de SV, el complejo sedimenta con un coeficiente de 8,2 S, el dímero de TcdA con un coeficiente de 3,8 S y el ARNt^{Lys}(UUU) con un coeficiente de 2,4 S. En el caso de los experimentos de SE, al calcular la masa flotante en cada uno de los casos se obtiene para el dímero de TcdA 63.531 \pm 840 Da (valor teórico 57,9 kDa), para el ARNt 26.066 \pm 156 Da (valor teórico 23,4 kDa) y para el complejo 114.297 \pm 824 Da (valor teórico 104,8 kDa). Estos datos encajan con la formación de un complejo formado por TcdA y dos moléculas de ARNt (tabla 5.2).

	MW teórico (Da)	MW experimental (Da)
TcdA (dímero)	57.900	63.531 \pm 840
tRNA^{Lys}(UUU)	23.400	26.066 \pm 156
TcdA-tRNA	104.800	114.297 \pm 824

Tabla 5.2. Masas moléculas teóricas y experimentales calculadas mediante SE-AUC para TcdA, tRNA^{Lys}(UUU) y el complejo formado entre estos dos componentes. Los pesos moleculares obtenidos experimentalmente encajan con los valores esperados en todos los casos, validando el experimento.

En cuanto a los ensayos del complejo mediante la técnica SAXS, ésta fue empleada con el fin de dilucidar la estructura tridimensional del complejo formado por TcdA y ARNt^{Lys}(UUU). Al emplear componentes previamente purificados para la formación del complejo, se colectaron datos tanto del complejo como del ARN sobrante en la mezcla realizada. Los datos de dispersión obtenidos fueron procesados e incorporados

Discusión

juntos en base a un radio de giro (Rg) coherente con los datos obtenidos. Se obtuvieron los siguientes datos: Rg del complejo 33,2 Å (puntos de Guinier 24-53; sRg 0,471-0,930; fidelidad 0,706), Rg 23,08 Å del ARNt libre (puntos de Guinier 5-72; sRg 0,139-0,892; fidelidad 0,979). De esta forma, tanto los datos de AUC (peso molecular del complejo 114 kDa) como los de SAXS (peso molecular del complejo 90 kDa) apoyan la existencia de un complejo donde se dan el dímero de TcdA junto con dos moléculas de ARN, al compararlo con los valores teóricos (peso molecular del complejo 104 kDa) (tabla 5.3). En la tabla 5.3 se comparan los diferentes valores obtenidos para los componentes del complejo TcdA-ARNt, demostrando la formación del dímero de proteína, unido a dos moléculas de ARNt, y por lo tanto la estequiometría 1:1.

	Masa teórica ¹ (Da)	SAXS (Guinier) ²	SAXS (Porod) ³	SE- AUC ⁴
Monómero TcdA	28.430,8	n.a.	n.a.	n.a.
Dímero TcdA	57.969,2	62.333,0	66.315,0	63.531,0
ARNt^{Lys}(UUU)	23.440,3	24.221,0	20.112,5	26.066,0
TcdA-ARNt	104.849,5	90.035,0 ⁵	148.805,0	114.297,0

Tabla 5.3. Comparación de las masas moleculares estimadas a partir de los datos de SAXS y SE-AUC del complejo TcdA-ARNt^{Lys}(UUU). En todos los casos las masas estimadas coinciden con los datos obtenidos mediante experimentos diferentes, corroborando la formación del complejo, TcdA-ARNt.

¹Las masas teóricas fueron calculadas usando la secuencia primaria menos la metionina inicial, la cual está ausente en la proteína recombinante TcdA. Las masas de 2 moléculas de ATP⁴⁻ (2 x 503,18 Da), 2 K⁺ (2 x 39,1 Da) y una de Na⁺ (23,0 Da) fueron añadidas al dímero de TcdA y a TcdA-ARNt. En la masa teórica para el ARNt^{Lys}(UUU) no se tomó en cuenta las modificaciones químicas.

²Masa molecular estimada derivada de la aproximación de Guinier.

³Masa molecular estimada del volumen de Porod.

⁴Masa molecular estimada calculada a partir de los datos de SE-AUC

La restauración de la forma del complejo *ab initio* (a partir de primeros principios) mediante los datos de SAXS se emplearon para generar una envoltura del complejo, así como del ARNt libre. De esta forma, se sugiere que el complejo está formado en su zona central por el dímero de TcdA, localizándose simétricamente a cada lado de éste una molécula de ARNt.

La reconstrucción del modelo del complejo mediante SAXS, se logró gracias al empleo de la estructura de TcdA resuelta en complejo con ATP mediante cristalografía de rayos-X y un ARNt modelo (pdb 4jxx). De nuevo se observa como la conformación aceptada anteriormente encaja con los cálculos empleando el ARNt modelo, es decir, con la molécula de proteína como parte central del complejo, donde se une una molécula de ARNt a cada lado de éste de forma simétrica. Es en esta orientación donde mayor implicación tienen las cargas positivas superficiales características del dímero de TcdA. Se observa como la zona plana debajo de la superficie de dimerización, así como el motivo de unión a K^+ , juegan un papel crucial en la unión al ARNt. Esta interfaz aplanada adopta la forma de una plataforma que facilita el acoplamiento del lado plano de la molécula de ARNt, mientras guiaría el anticodón del ácido nucleico (zona anexa a la base que es capaz de modificar TcdA, la base A37 del ARNt) hacia el centro activo de la proteína, donde se encuentra la Cys66. Las propiedades electrostáticas tanto de la proteína como del ARNt aportan la estabilidad necesaria para la formación del complejo. Esta zona con cargas positiva en TcdA, incluye la zona de unión a ATP (o AMP), lo que provee una extensa zona cargada, facilitando la unión del ARNt a través de las cargas negativas de los fosfatos (figura 5.6). Por lo tanto fue posible describir una nueva forma de interacción proteína-ARNt en el caso de TcdA, hecho novedoso.

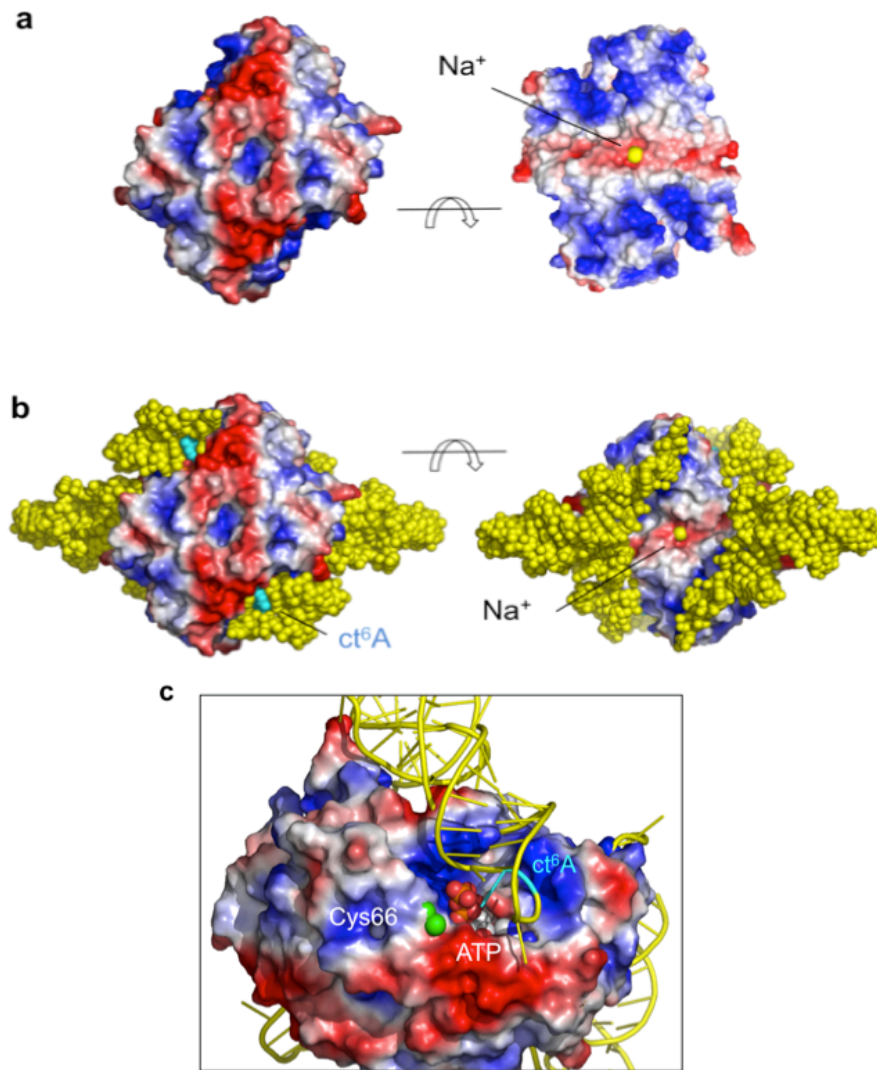


Figura 5.6. Modelo de interacción TcdA-ARNt. **a** Superficie electrostáticamente calculada con APBS (*Adaptive Poisson-Boltzmann Solver*) (Baker et al., 2001) y modelado finalmente con PyMOL (Schrodinger, 2010) del dímero de TcdA. Se muestran dos vistas rotadas una respecto de la otra 90° en torno al eje horizontal. El catión Na⁺ se muestra como una esfera amarilla. **b** Complejo TcdA-ARNt. El ARNt se modeló en base al resultado obtenido de los experimentos de SAXS del complejo TcdA-ARNt^{Lys}(UUU), donde la mayoría de la superficie está cargada positivamente. **c** Detalle de la vista de la interacción entre TcdA-ARNt. La proteína se representa como en el caso de (a), con el grupo tiol de Cys66 en verde. La molécula de ARNt introduce la zona del anticodón en el centro activo de TcdA, donde se encuentra la molécula de ATP (en esferas y con colores CPK), con ct⁶A37 en cian, encarando el centro activo (López-Esteva *et al.*, 2015).

En la figura 5.7 se muestra un esquema general de la relación entre el sistema CSD de *E.coli* en cuanto a la transferencia de átomos de S, y la modificación de los ARNt que reconoce codones de tipo ANN por parte de la proteína TcdA. Estos dos sistemas fueron relacionados en trabajos anteriores (Trotter *et al.*, 2009; Miyauchi *et al.*, 2013). Sin embargo, hasta ahora no se conocían el modo de interacción a nivel molecular lo que muestra con más detalle el mecanismo entre los dos sistemas y permite entenderlo mejor.

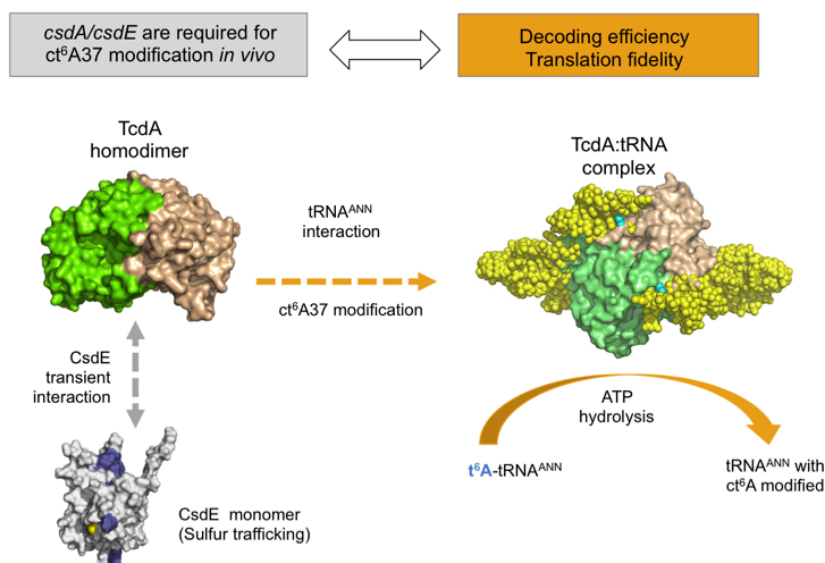


Figura 5.7. Relación entre el sistema CSD de *E. coli* y la modificación de ct^6A37 ($ARNt^{ANN}$). Se representa la superficie de TcdA (uno de los monómeros en verde y el otro en color trigo) interactuando de forma transitoria pero específica con el aceptor de azufre del sistema CSD, CsdE (en gris, con la superficie de interacción con TcdA en azul y con la Cys61 en amarillo), conectando de esta manera con el sistema de transferencia de azufre mencionado. La interacción transitoria TcdA-CsdE se representa con la flecha gris en doble dirección. TcdA interactúa también con $ARNt$ capaces de reconocer codones de tipo ANN. Esta interacción resulta en un complejo con estequiometría 1:1, donde $TcdA_2$ es el núcleo del complejo y gracias a la formación del cual se puede llevar a cabo la hipermodificación de la base t^6A37 de este tipo de $ARNts$ (López-Esteva *et al.*, 2015).

5.7. Estructura de la PRPP sintetasa (PrsA) de *E. coli*

PrsA, o PRPP sintetasa es una enzima clave en varias rutas biosintéticas. PrsA cataliza la formación de PRPP (y AMP) a partir de ribosa 5-pentosa (R5P) y ATP,

permitiendo la comunicación del flujo metabólico entre la ruta de las pentosas fosfato y las rutas de recuperación y de síntesis *de novo* de nucleótidos de pirimidinas y purinas; además, las PRPP sintetetasas actúan en las rutas de biosíntesis de la histidina, el triptófano y las coenzimas de nucleótidos de piridina. Dada la posición central del PRPP en el metabolismo celular, la función de las PRPP sintetetasas es esencial para la vida. Se han realizado muchos estudios detallados sobre su mecanismo enzimático, cinética, relación estructura-función, papel del fosfato inorgánico y Mg^{2+} como activadores, su comportamiento alostérico y su arquitectura cuaternaria.

En este trabajo nos interesaba entender la estructura de PrsA desde dos puntos de vista. El primero, como posible interactor del sistema CSD, y el segundo, como posible diana para el descubrimiento de nuevos antibióticos.

El primer estudio donde se describió la posibilidad de que PrsA interaccionase con el sistema CSD también identificó CsdL/TcdA como interactor de CsdE (Trotter *et al.*, 2009). Si bien la identificación de la interacción CsdL/TcdA no fue sorprendente y ha podido ser corroborada posteriormente (López-Esteva *et al.*, 2015), la interacción de PrsA/CsdE aún permanece sin caracterizar. A pesar de haber realizado numerosos esfuerzos para capturar una posible interacción entre PrsA y CsdE, todavía no hemos sido capaces de confirmarlo; ni siquiera técnicas capaces de detectar interacciones de baja afinidad y/o transitorias han permitido detectarla. Por tanto, hasta este momento, debemos concluir que dicha interacción o no existe en condiciones fisiológicas o es extremadamente débil.

En cuanto a su uso potencial como diana de nuevos fármacos antibacterianos, la determinación de la estructura de PrsA en complejo con ATP a alta resolución (1,77 Å) abre la posibilidad de comprender en detalle atómico los determinantes estructurales de la interacción de PrsA con sus sustratos y de iniciar búsquedas sistemáticas de inhibidores potenciales, tanto experimentalmente como *in silico*. La única estructura de EcPrsA conocida hasta ahora tiene el sitio activo vacío (pdb 4s2u), por lo que no aporta los detalles precisos de la interacción que sí pueden obtenerse experimentalmente de nuestra estructura.

En comparación con estructuras de PRPP sintetetasas, la arquitectura global de PrsA está muy conservada. Por ejemplo, EcPrsA puede superponerse con la PRPP sintetasa de *B. subtilis* (pdb 1dku) con r.m.s.d. de 1,16 Å (296 C- α). Las mayores diferencias se encuentran en el esqueleto polipeptídico de las regiones más flexibles de la estructura (*e.g.*, la lámina β perpendicular a la central, los lazos más expuestos al solvente) así como en aquellas zonas que no son visibles en algunas estructuras por ausencia de suficiente orden estructural. En cuanto a las cadenas laterales de los residuos

conservados del centro activo, los diferentes rotámeros que se observan tienden a correlacionar con la presencia o ausencia de los ligandos, sustratos y análogos con los que se han cristalizado las PRPP sintetasas.

En PRPP sintetasas se ha descrito un segundo sitio de unión a nucleótido que serviría para unir una molécula de ADP con un papel regulador, que tendría implicaciones para el comportamiento alostérico de PrsA. Los residuos descritos en este sitio regulador pertenecen a tres subunidades diferentes, en un bolsillo cercano al sitio de ATP del sitio activo. En *EcPrsA*, los residuos correspondientes están conservados, e incluye Gln162A, Asp170A, Asn171A, y el C-terminal (Ser334A-Phe339A) de una subunidad (A), los residuos Arg126B-Arg131B del lazo flexible de una segunda subunidad (B), y los cuatro residuos de una tercera subunidad (C), Asn73C, Arg75C, Ala106C, y Ser107C.

La conservación general de los residuos del sitio catalítico y del sitio regulador para el ADP (el sitio para el fosfato inorgánico, Pi, que también tiene un papel regulador/activador, también está conservado) sugieren que los mecanismos enzimáticos de activación, regulación y catálisis están bien conservados a través de muchas enzimas bacterianas. Esto, junto con el papel esencial de PrsA en el metabolismo celular bacteriano, de aminoácidos y de nucleótidos, permiten identificarla como una diana farmacéutica potencial. La disponibilidad de la estructura de PrsA de *E. coli* a 1,77 Å en complejo con ATP añade información estructural importante al conocimiento que se tiene hasta ahora de las PRPP sintetasas y podrá ser empleado para el descubrimiento de nuevas moléculas con capacidad inhibitoria.

5.8. Estructura de *EcOPRT* en complejo con 5-FOA y 5-AOA

La estructura de *EcOPRT* ha sido resuelta por cristalografía de rayos-X en presencia de los análogos estructurales de orotato 5'-FOA (5-fluoro-orotato) y 5'-AOA (5-amino-orotato). En ambos casos la resolución es muy alta, < 2.0 Å, lo que permite definir con gran precisión la estructura de los análogos y las distancias entre ellas y los residuos del sitio activo de *EcOPRT* con los que interactúan. La semejanza química entre ambas moléculas, que difieren exclusivamente en la identidad del sustituyente en la posición 5' del anillo de orotato (F en 5'-FOA y NH₂ en 5'-AOA), sugieren que el sitio de unión será el mismo en ambos casos. De hecho, 5'-FOA y 5'-AOA interaccionan con los mismos residuos de *EcOPRT*, la cadena principal de Phe34, Phe35, Lys26, y la cadena lateral de Arg156. A pesar de la similitud química y la identidad en el lugar y modo de interacción con *EcOPRT*, solamente 5'-FOA es capaz de reaccionar, dando lugar a 5-

fluoro-orotidina 5'-fosfato como producto de reacción. Este producto es metabolizado hasta 5-fluorouracil, inhibidor de la timidato sintasa (Ong *et al.*, 2013). Es por esto que 5'-FOA se ha denominado una prodroga, ya que, tras transformarse en 5-fluoro-orotidina 5'-fosfato, puede actuar inhibiendo una ruta vital para el metabolismo celular de ácidos nucleicos. Al contrario que 5'-FOA, 5'-AOA no sufre conversión o lo hace muy lentamente, de manera que se asemeja más en su comportamiento cinético a un inhibidor competitivo con orotato. La radical diferencia en cuanto al comportamiento como sustratos de ambas moléculas se encuentra en la reactividad química y enzimática del orotato sustituido; el átomo de F en 5'-FOA es más electronegativo que el heterociclo del orotato, mientras que el grupo amino primario del 5'-AOA lo es menos, de manera que el análogo de sustrato es activado en el caso del 5'-FOA e inactivado en el caso del 5'-AOA con respecto al ataque por la *EcOPRT*.

En cuanto a las comparaciones estructurales con la proteína *apo* resuelta anteriormente (pdb 1oro), no existen diferencias en cuanto a la estructura, respetándose todos los componentes estructurales. En la figura 5.8 se muestra una superposición de la proteína *apo*, tanto con la estructura obtenida con 5-FOA, como con la obtenida con 5-AOA, donde se respeta la estructura global del dímero.

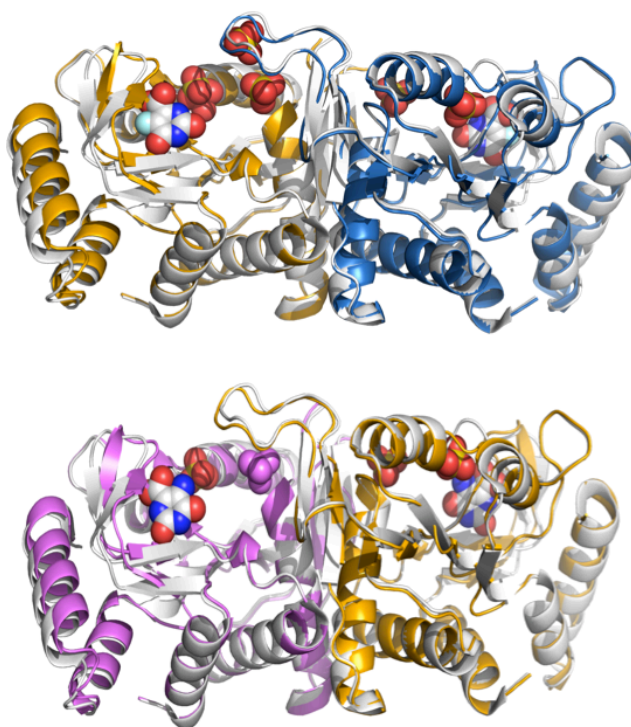


Figura 5.8. Superposición de la OPRT/5-FOA (arriba) y OPRT/5-AOA (abajo) con la enzima *apo*. En ambos casos, los análogos de orotato (5-FOA y 5-AOA) están representados en modelos de esferas y colores CPK. La estructura de referencia (OPRT *apo*) se muestra en color gris claro.

5.9. Comparación estructural de *MtOPRT* y *EcOPRT*

La estructura cristalográfica de *MtOPRT* a 1,54 Å de resolución es la primera estructura resuelta de esta enzima. El hecho de que enzimas homólogas a OPRT de otros patógenos se hayan utilizado como dianas farmacológicas y la enorme trascendencia de *M. tuberculosis* como una amenaza a la salud pública, justifican nuestro interés en obtener una estructura de muy alta calidad y resolución. Si bien la arquitectura de *MtOPRT* es muy semejante a la de *EcOPRT* y, de hecho, al resto de PRTasas de tipo I resueltas hasta ahora, hay ciertas diferencias que podrían resultar interesantes desde el punto de vista del desarrollo de nuevos fármacos. La superposición de las estructuras de *MtOPRT* y *EcOPRT* como monómeros tiene un r.m.s.d de 3,63 Å y demuestra diferencias significativas entre ambas estructuras que se comentan a continuación:

i) La ausencia casi total del motivo C-terminal consistente en dos hélices α , que justifica la longitud menor de *MtOPRT*, y que en *EcOPRT* empaqueta directamente contra los motivos N-terminales (la hélice α 1 y la “tapa”); el lazo 157-165 que en *EcOPRT* interacciona con las hélices α terminales y con la “tapa”, en *MtOPRT* es un giro corto formado por 3 residuos únicamente (152-154). En *MtOPRT*, la ausencia de la estabilización proporcionada por las dos hélices α C-terminales y el giro más corto provoca que la “tapa”, que es esencial para la actividad enzimática y que une 5'-FOA/5'-AOA en *EcOPRT*, se encuentre en una conformación más abierta, más expuesta al solvente. Además, las hélices α 2 y 3, que siguen a la tapa, son más cortas y están unidas mediante un lazo más corto que en *EcOPRT*. Estos factores convierten *MtOPRT* en una enzima mucho más compacta. En cuanto a la “tapa”, su composición y longitud es muy similar entre ambas enzimas; la principal diferencia entre ambas estructuras es la posición y orientación de las dos hebras β de la punta de la “tapa” (His20-Tyr33), que están más inclinadas hacia el solvente que en *EcOPRT*. De hecho, existe un desplazamiento de hasta 6 Å entre la posición de la “tapa” en *MtOPRT* con respecto a *EcOPRT*, posiblemente debido a la ausencia en *MtOPRT* del lazo largo descrito anteriormente.

ii) Los dímeros de *MtOPRT* y *EcOPRT* son mucho más difíciles de interpretar porque la arquitectura de los dos dímeros, aunque coincide en la topología y estructura general, diverge notablemente en las dimensiones. La superposición de los dímeros de *MtOPRT* y *EcOPRT* no es realmente posible ya que las diferencias estructurales se amplifican a medida que consideramos regiones más alejadas de la interfaz del dímero; las superposiciones como monómeros tienen un r.m.s.d. con valor de que 3,63 Å. De esta forma, al superponer una sola subunidad del dímero la otra queda desplazada

completamente, imposibilitando la comparación. Por ejemplo, la lámina β central de la subunidad no superimpuesta queda desplazada por una hebra β completa. La interfaz y estructura del dímero de *MtOPRT* es más compacta que en *EcOPRT*.

En cualquier caso, el lazo catalítico de *MtOPRT* presenta una region invisible en densidad electrónica, entre los residuos 97-103, que presumiblemente se insertarán en el centro activo de la subunidad vecina durante la catálisis. La estructura de *MtOPRT* presta soporte a esta idea ya que la distancia entre este lazo desordenado y el sitio activo vecino es compatible con esta función del lazo catalítico. En la figura 5.9 se muestra la superposición de las dos estructuras.

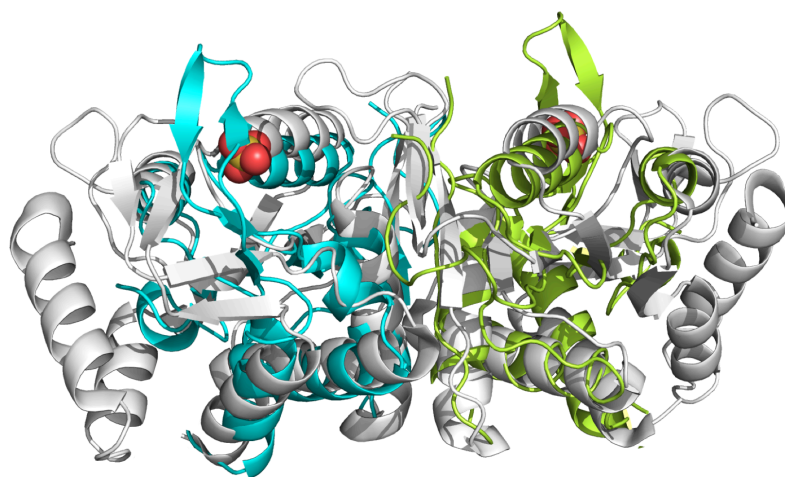


Figura 5.9. Superposición de *MtOPRT* y *EcOPRT*. El dímero de *MtOPRT* está coloreado por cadenas (cián y verde), mientras que *EcOPRT* se muestra en gris claro. Los iones fosfato se representan como modelos de esferas. Es de destacar la ausencia de las dos hélices α C-terminales en el caso de *MtOPRT*. En consecuencia, el lazo de *EcOPRT* que se empaquetaba contra ellas ha sido reducido en *MtOPRT* a un giro mucho más corto (152-155 en *MtOPRT* frente a 157-165 en *EcOPRT*).

5.10. *EcOPRT* y *MtOPRT* son capaces de metabolizar el análogo de OA, 5-FOA

En ambas enzimas, tanto la de *E. coli* como la de *M. tuberculosis*, se observa la capacidad para metabolizar el ácido orótico y el análogo 5-FOA. En el caso de *E. coli*, se muestra una menor afinidad por el sustrato natural ($K_m = 94 \mu M$) que en otros casos descritos, como es en la enzima de *S. typhimurium*, mostrando ésta tres veces mayor

afinidad ($K_m = 28 \mu\text{M}$) por el mismo sustrato (Bhatia *et al.*, 1992), aunque los valores de k_{cat} son idénticos ($k_{\text{cat}} = 27 \text{ s}^{-1}$). En cuanto a la capacidad que tiene la enzima de *E. coli* de metabolizar el análogo 5-FOA, se observa que es capaz pero con una menor afinidad hacia este sustrato ($K_m = 171 \mu\text{M}$) que hacia el OA. Sin embargo, la k_{cat} es el doble ($k_{\text{cat}} = 68 \text{ s}^{-1}$) para el análogo, tanto en el caso de *E. coli* como en el de *S. typhimurium*. De esta forma, aunque 5-FOA no sea un inhibidor directo de esta enzima si se ha descrito como inhibidor de la enzima timidilato sintasa, tras metabolizarse a 5-FUMP, en el organismo *T. brucei* (Ong *et al.*, 2013), bloqueando la síntesis de pirimidinas. Además, la relación k_{cat}/K_m es superior en el caso de 5-FOA, por lo que podría llegar a competir por el sustrato natural e inhibir parte de las funciones celulares. Se realizaron ensayos de crecimiento de *E. coli* (XL1blue) en LB, donde la adición de 5-FOA al medio parece ralentizar el incremento de la D.O.₅₉₀, a 310 K (los datos no se muestran), en las primeras fases del crecimiento de masa.

En el caso de la enzima de *M. tuberculosis* ocurre algo similar, aunque ha sido imposible cristalizar la enzima con 5-FOA. En este caso la afinidad por ambos sustratos es similar, observándose una $K_m = 10 \mu\text{M}$ en el caso del OA (Breda *et al.*, 2012) y una $K_m = 9,55 \mu\text{M}$ en el caso de 5-FOA. Sin embargo, al calcular la k_{cat} ésta es casi siete veces mayor en el caso de 5-FOA ($4,2 \text{ s}^{-1}$), lo que convierte a este análogo en un posible inhibidor de la ruta de síntesis de timidinas en este patógeno y por lo tanto en un potencial antibiótico, ya que es capaz de metabolizarse más rápido por parte de la OPRT.

Conclusiones

6. Conclusiones

1. La estructura del complejo (CsdA-CsdE)₂ persulfurado, resuelta por cristalografía de rayos-X a 2,49 Å, permite entender el papel estructural y funcional de la secuencia motivo conservada característica de las cisteína desulfurasas de tipo II, que proporciona estabilización electrostática al grupo persulfuro durante su transferencia desde el donador de azufre (CsdA) al aceptor de azufre (CsdE).
2. Las estructuras tridimensionales de CsdA en estado apo (no modificado), persulfurada, S-sulfatada (oxidada), y de los mutantes H120A y R250A, han sido resueltas por cristalografía de rayos-X. La estructura de CsdA oxidada proporciona una explicación a nivel atómico de la no funcionalidad del sistema CSD bajo estrés oxidativo.
3. En base a la homología estructural y de secuencia entre, respectivamente, CsdA y CsdE del sistema CSD y SufS y SufE del sistema Suf, se ha demostrado que CsdA es capaz de transferir átomos de azufre a SufE, *i.e.*, que existe “*crosstalk*” entre los sistemas CSD y Suf.
4. Las estructuras cristalográficas de TcdA (TcdA, la N⁶-treonilcarbamoiladenosina deshidratasa responsable de la ciclación de t⁶A37 en los ARNt^{ANN}) con ATP y con AMP a 1,77 Å y 1,80 Å de resolución, respectivamente han mostrado modificaciones funcionalmente relevantes del plan estructural de la superfamilia de enzimas E1. Además de ser dimérica, TcdA presenta un intercambio de motivos estructurales que altera la topología canónica, la generación de un motivo de unión a K⁺ que sustituye el motivo de unión a Zn²⁺ clásico, y la presencia de superficie cargada positivamente. Como consecuencia muestra una actividad deshidratasa.
5. TcdA interacciona con ARNt^{Lys}(UUU) con estequiometría 1:1. En base a difracción de rayos-X de bajo ángulo, además del empleo de técnicas bioquímicas/biofísicas (AUC, EMSA), se ha construido un modelo híbrido del complejo TcdA-ARNt con el ASL del ARNt sustrato en el surco del centro activo de TcdA, y la cadena lateral del ribonucleósido t⁶A37 directamente enfrentada a los grupos catalíticos de TcdA y de ATP.
6. TcdA interacciona de forma directa con el sistema CSD, a través de los dos componentes de éste, CsdA y CsdE, siendo más fuerte la interacción TcdA-CsdE, que forman un complejo con estequiometría 1:1. El complejo TcdA-CsdE pudo ser reconstruido estructuralmente mediante la combinación de entrecruzamiento con

Conclusiones

BMOE con el análisis del complejo TcdA-CsdE por difracción de rayos-X de bajo ángulo.

7. Se ha cristalizado la PRPP sintetasa (PrsA) de *E. coli* en presencia de ATP, un posible interactor de CsdE. La calidad de la difracción que puede obtenerse de estos cristales alcanza 1,77 Å y por tanto hace posible el estudio a nivel atómico de esta enzima clave en la ruta de síntesis de las pirimidinas.
8. La estructura cristalográfica de la orotato fosforribosil transferasa OPRT de *E. coli* ha sido resuelta en complejo con dos análogos de sustrato (5-FOA y 5-AOA), a resoluciones de 1,43 y 1,62 Å, respectivamente. Las estructuras revelan el modo de unión de estos compuestos a la OPRT. Ambos compuestos actúan como prodrogas, siendo los productos de la reacción inhibidores del metabolismo de pirimidinas.
9. La estructura cristalográfica de la OPRT de *M. tuberculosis* ha sido resuelta por primera vez, a 1,54 Å de resolución. La estructura, en estado apo, revela diferencias interesantes con la enzima homóloga de *E. coli*. La OPRT de *M. tuberculosis* podría representar una nueva diana en la búsqueda de medicinas contra la tuberculosis.
10. La caracterización cinética de OPRT de *E. coli* y de *M. tuberculosis* con respecto a sus sustratos naturales y a las prodrogas 5-FOA y 5-AOA indica que ambas enzimas son capaces de metabolizar más rápidamente la prodroga 5-FOA que los propios sustratos naturales.

Conclusions

7. Conclusions

1. The crystal structure of the persulfurated (CsdA-CsdE)₂ complex, resolved at 2.49 Å, has unveiled the catalytic mechanism of the transfer of S atoms across protein-protein interfaces during the transpersulfuration reaction in the CSD system. The structure has also revealed the structural and functional role of the conserved motif sequence of the type II cysteine desulfurases, which provides electrostatic stabilization to the persulfide group during transfer from CsdA to CsdE.
2. The structures of CsdA in the apo, persulfurated and S-sulfated forms, and CsdA mutants H120A and R250A, have been solved by X-ray crystallography. The oxidized form of CsdA provides an atomic-level explanation of the impairment of the CSD system under oxidative conditions.
3. CsdA can transfer S atoms to the sulphur acceptor of the Suf system, thereby demonstrating the possibility of crosstalk between the CSD and Suf systems.
4. The crystallographic structures of *N*⁶-threonylcarbamoyl dehydratase, TcdA, from *E. coli* with ATP and AMP at 1.77 and 1.80 Å resolution, respectively, reveal unique features in the E1-like superfamily. Besides being a homodimer, TcdA presents a topological swap between the metal binding motif and the C-terminal helix. Also, TcdA is the first protein in the E1-like superfamily that binds K⁺ and Na⁺, instead of Zn²⁺, displays a new activity, and is able to bind tRNA.
5. TcdA interacts with tRNA^{Lys}(UUU) with a 1:1 stoichiometry. A hybrid model based on SAXS, biochemical and biophysical techniques (EMSA, AUC) has been generated of the TcdA-tRNA complex. This model shows the anti-codon stem loop of tRNA, with the substrate t⁶A37, reaching into the TcdA active site, in a conformation compatible with the catalysis.

Conclusions

6. TcdA physically interacts with CsdE and, much less strongly, with CsdA. A structural model of the TcdA-CsdE has been built based on SAXS shape reconstruction of the heterodimeric complex crosslinked with BMOE.
7. The crystal structure of PRPP synthetase from *E. coli*, a possible interactor of CsdE, has been resolved at 1.77 Å resolution with ATP bound in the active site.
8. The crystal structures of OPRT from *E. coli* in complex with the substrate analogues 5-FOA y 5-AOA have been determined at 1.43 and 1.62 Å resolution, respectively. Both molecules act as prodrugs that can block the route of pyrimidine synthesis.
9. The crystal structure of OPRT from the human pathogen *M. tuberculosis* has been determined at 1.54 Å in its apo form. The structure of MtOPRT reveals unique features that can be exploited for the development of novel antibiotics.
10. The kinetic characterization of *E. coli* and *M. tuberculosis* OPRT with their natural substrates and the 5-FOA and 5-AOA prodrugs indicate that both enzymes can metabolize 5-FOA more efficiently than the natural substrates.

Referencias

8. Referencias

- Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*. 2012; 68: 352–367. doi: 10.1107/S0907444912001308 PMID: 22505256
- Alexander FW, Sandmeier E, Mehta PK, Christen P. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific α , β and γ families. *Eur J Biochem*. 1994 Feb 1;219(3):953-60.
- Alfonzo JD, Lukeš J. Assembling Fe/S-clusters and modifying tRNAs: ancient co-factors meet ancient adaptors. *Trends Parasitol*. 2011 Jun;27(6):235-8. doi: 10.1016/j.pt.2011.02.003. Epub 2011 Mar 17.
- Ayala-Castro C, Saini A, Outten FW. Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev*. 2008 Mar;72(1):110-25, table of contents. doi: 10.1128/MMBR.00034-07.
- Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A*. 2001 Aug 28;98(18):10037-41. Epub 2001 Aug 21.
- Barras F, Loiseau L, Py B. How *Escherichia coli* and *Saccharomyces cerevisiae* build Fe/S proteins. *Adv Microb Physiol*. 2005;50:41-101.
- Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr*. 2011 Apr;67(Pt 4):271-81. doi: 10.1107/S0907444910048675. Epub 2011 Mar 18.
- Baugh L, Gallagher LA, Patrapuvich R, Clifton MC, Gardberg AS, Edwards TE, Armour B, Begley DW, Dieterich SH, Dranow DM, Abendroth J, Fairman JW, Fox D 3rd, Staker BL, Phan I, Gillespie A, Choi R, Nakazawa-Hewitt S, Nguyen MT, Napuli A, Barrett L, Buchko GW, Stacy R, Myler PJ, Stewart LJ, Manoil C, Van Voorhis WC. Combining functional and structural genomics to sample the essential *Burkholderia* structome. *PLoS One*. 2013;8(1):e53851. doi: 10.1371/journal.pone.0053851. Epub 2013 Jan 31.
- Berg JM, Tymoczko JL, Stryer L. *Biochemistry* 5th Edition. New York: W H Freeman; 2002.
- Bhatia MB, Grubmeyer C. The role of divalent magnesium in activating the reaction catalyzed by orotate phosphoribosyltransferase. *Arch Biochem Biophys*. 1993 Jun;303(2):321-5.
- Björk G.R. Biosynthesis and function of modified nucleosides. In Soll,D. and RajBhandary,U.L. (eds), *tRNA: Structure, Biosynthesis and Function*. ASM, Washington DC, 1995 pp. 165–205.
- Block H, Maertens B, Spriestersbach A, Brinker N, Kubicek J, Fabis R, Labahn J, Schäfer F. Immobilized-metal affinity chromatography (IMAC): a review. *Methods Enzymol*. 2009;463:439-73. doi: 10.1016/S0076-6879(09)63027-5.
- Bolstad HM, Botelho DJ, Wood MJ. Proteomic analysis of protein-protein interactions within the Cysteine Sulfinate Desulfinate Fe-S cluster biogenesis system. *J Proteome Res*. 2010 Oct 1;9(10):5358-69. doi: 10.1021/pr1006087.

Referencias

- Bower SG, Harlow KW, Switzer RL, Hove-Jensen B. Characterization of the *Escherichia coli* prsA1-encoded mutant phosphoribosylpyrophosphate synthetase identifies a divalent cation-nucleotide binding site. *J Biol Chem*. 1989 Jun 15;264(17):10287-91.
- Breda A, Martinelli LK, Bizarro CV, Rosado LA, Borges CB, Santos DS, Basso LA. Wild-type phosphoribosylpyrophosphate synthase (PRS) from *Mycobacterium tuberculosis*: a bacterial class II PRS? *PLoS One*. 2012;7(6):e39245. doi: 10.1371/journal.pone.0039245. Epub 2012 Jun 20.
- Burgess RR. Protein Precipitation Techniques. *Methods Enzymol*. 2009;463:331-42. doi: 10.1016/S0076-6879(09)63020-2.
- Burroughs AM, Iyer LM, Aravind L. Natural history of the E1-like superfamily: implication for adenylation, sulfur transfer, and ubiquitin conjugation. *Proteins*. 2009 Jun;75(4):895-910. doi: 10.1002/prot.22298.
- Chahal HK, Dai Y, Saini A, Ayala-Castro C, Outten FW. The SufBCD Fe-S scaffold complex interacts with SufA for Fe-S cluster transfer. *Biochemistry*. 2009 Nov 10;48(44):10644-53. doi: 10.1021/bi901518y.
- Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010; 66: 12-21. doi: 10.1107/S0907444909042073 PMID: 20057044
- Cherney MM, Cherney LT, Garen CR, James MN. The structures of *Thermoplasma volcanium* phosphoribosyl pyrophosphate synthase bound to ribose-5-phosphate and ATP analogs. *J Mol Biol*. 2011 Nov 4;413(4):844-56. doi: 10.1016/j.jmb.2011.09.007. Epub 2011 Sep 21.
- Cole JL, Lary JW, P Moody T, Laue TM. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods Cell Biol*. 2008;84:143-79.
- Cole JL. Analysis of heterogeneous interactions. *Methods Enzymol*. 2004; 384: 212-232. PMID: 15081689
- Cunningham JT, Moreno MV, Lodi A, Ronen SM, Ruggero D. Protein and nucleotide biosynthesis are coupled through a single rate limiting enzyme, PRPS2, to drive cancer. *Cell*. 2014 May 22;157(5):1088-103. doi: 10.1016/j.cell.2014.03.052.
- Dai Y, Outten FW. The *E. coli* SufS-SufE sulfur transfer system is more resistant to oxidative stress than IscS-IscU. *FEBS Lett*. 2012 Nov 16;586(22):4016-22. doi: 10.1016/j.febslet.2012.10.001. Epub 2012 Oct 12.
- Di Salvo ML, Budisa N and Contestabile R. PLP-dependent enzymes a powerful tool for metabolic synthesis of non-canonical aminoacids. Beilstein Bozen Symposium on Molecular Engineering and Control May 14th – 18th, 2012, Prien (Chiemsee), Germany.
- Dobbek H, Svetlitchnyi V, Gremer L, Huber R, Meyer O (2001) Crystal structure of a carbon monoxide dehydrogenase reveals a [Ni-4Fe-5S] cluster. *Science*. 2001 Aug 17;293(5533):1281-5.
- Dodson E (2003) Is it jolly SAD? *Acta Crystallogr D Biol Crystallogr*. Nov;59 (Pt 11):1958-65. Epub 2003 Oct 23.

- Drenth, J. Crystallizing a protein. Principles of Protein X-Ray Crystallography (chapter I). 2007 463:1-20. doi: 10.1007/0-387-33746-6.
- Drenth, J. Molecular Replacement (chapter X). Principles of Protein X-Ray Crystallography. 2007 463:210-230. doi: 10.1007/0-387-33746-6.
- Drenth, J. Refinement of the model structure (chapter XIII). Principles of Protein X-Ray Crystallography. 2007 463:248-278. doi: 10.1007/0-387-33746-6.
- Du Q, Wang H, Xie J. Thiamin (vitamin B1) biosynthesis and regulation: a rich source of antimicrobial drug targets? Int J Biol Sci. 2011 Jan 9;7(1):41-52.
- Duda DM, Walden H, Sfondouris J, Schulman BA. Structural analysis of Escherichia coli ThiF. J Mol Biol. 2005 Jun 17;349(4):774-86.
- El Yacoubi B, Bailly M, de Crécy-Lagard V. Biosynthesis and function of posttranscriptional modifications of transfer RNAs. Annu Rev Genet. 2012;46:69-95. doi: 10.1146/annurev-genet-110711-155641. Epub 2012 Aug 16.
- Eliot AC, Kirsch JF. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. Annu Rev Biochem. 2004;73:383-415.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010; 66: 486–501. doi: 10.1107/S0907444910007493 PMID: 20383002
- Eriksen TA, Kadziola A, Larsen S. Binding of cations in Bacillus subtilis phosphoribosyldiphosphate synthetase and their role in catalysis. Protein Sci. 2002 Feb;11(2):271-9.
- Esberg B, Leung HC, Tsui HC, Björk GR, Winkler ME. Identification of the miaB gene, involved in methylthiolation of isopentenylated A37 derivatives in the tRNA of Salmonella typhimurium and Escherichia coli. J Bacteriol. 1999 Dec;181(23):7256-65.
- Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011 Apr;67(Pt 4):282-92. doi: 10.1107/S090744491003982X. Epub 2011 Mar 18.
- Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013 69: 1204–1214. doi: 10.1107/S0907444913000061 PMID: 23793146
- Fernandez FJ, de Vries D, Peña-Soler E, Coll M, Christen P, Gehring H, Vega MC. Structure and mechanism of a cysteine sulfinatase engineered on the aspartate aminotransferase scaffold. Biochim Biophys Acta. 2012 Feb;1824(2):339-49. doi: 10.1016/j.bbapap.2011.10.016. Epub 2011 Nov 23.
- Finehout EJ, Lee KH. An introduction to mass spectrometry Biochem Mol Biol Educ. 2004 Mar;32(2):93-100. doi: 10.1002/bmb.2004.494032020331.y applications in biological research.

Referencias

- Flint DH, Allen RM. Iron-sulfur proteins with nonredox functions. *Chem Rev.* 1996 Nov 7;96(7):2315-2334.
- Fox BA, Bzik DJ. De novo pyrimidine biosynthesis is required for virulence of *Toxoplasma gondii*. *Nature.* 2002 Feb 21;415(6874):926-9.
- Franke D, Svergun D.I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Cryst.* (2009). 42, 342-346 doi:10.1107/S0021889809000338
- Frazzon J, Fick JR, Dean DR. Biosynthesis of iron-sulphur clusters is a complex and highly conserved process. *Biochem Soc Trans.* 2002 Aug;30(4):680-5.
- Freude K, Hoffmann K, Jensen LR, Delatycki MB, des Portes V, Moser B, Hamel B, van Bokhoven H, Moraine C, Fryns JP, Chelly J, Géczy J, Lenzner S, Kalscheuer VM, Ropers HH. Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. *Am J Hum Genet.* 2004 Aug;75(2):305-9. Epub 2004 May 25.
- Fridman A, Saha A, Chan A, Casteel DE, Pilz RB, Boss GR. Cell cycle regulation of purine synthesis by phosphoribosyl pyrophosphoribosyl pyrophosphate and inorganic phosphate. *Biochem J.* 2013 Aug 15;454(1):91-9. doi: 10.1042/BJ20130153.
- Gao G, Williams JG, Campbell SL. Protein-protein interaction analysis by nuclear magnetic resonance spectroscopy. *Methods Mol Biol.* 2004;261:79-92.
- Gavira JA. Current trends in protein crystallization. *Arch Biochem Biophys.* 2015 Dec 31. pii: S0003-9861(15)30121-1. doi: 10.1016/j.abb.2015.12.010.
- Goldsmith-Fischman S, Kuzin A, Edstrom WC, Benach J, Shastry R, Xiao R, Acton TB, Honig B, Montelione GT, Hunt JF. The SufE sulfur-acceptor protein contains a conserved core structure that mediates interdomain interactions in a variety of redox protein complexes. *J Mol Biol.* 2004 Nov 19;344(2):549-65.
- González-Segura L, Witte JF, McClard RW, Hurley TD. Ternary complex formation and induced asymmetry in orotate phosphoribosyltransferase. *Biochemistry.* 2007 Dec 11;46(49):14075-86. Epub 2007 Nov 20.
- Grabski AC. Advances in preparation of biological extracts for protein purification. *Methods Enzymol.* 2009;463:285-303. doi: 10.1016/S0076-6879(09)63018-4.
- Grishin NV, Phillips MA, Goldsmith EJ. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci.* 1995 Jul;4(7):1291-304.
- Grubmeyer C, Hansen MR, Fedorov AA, Almo SC. Structure of *Salmonella typhimurium* OMP synthase in a complete substrate complex. *Biochemistry.* 2012 Jun 5;51(22):4397-405. Epub 2012 May 23.
- Gustilo EM, Vendeix FA, Agris PF. tRNA's modifications bring order to gene expression. *Curr Opin Microbiol.* 2008 Apr;11(2):134-40. doi: 10.1016/j.mib.2008.02.003. Epub 2008 Apr 2.
- Hassell AM, An G, Bledsoe RK, Bynum JM, Carter HL 3rd, Deng SJ, Gampe RT, Grisard TE, Madauss KP, Nolte RT, Rocque WJ, Wang L, Weaver KL, Williams SP, Wisely GB, Xu R,

- Shewchuk LM. Crystallization of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*. 2007 Jan;63(Pt 1):72-9. Epub 2006 Dec 13.
- Helm M, Alfonzo JD. Posttranscriptional RNA Modifications: playing metabolic games in a cell's chemical Legoland. *Chem Biol*. 2014 Feb 20;21(2):174-85. doi: 10.1016/j.chembiol.2013.10.015. Epub 2013 Dec 5.
- Henriksen A, Aghajari N, Jensen KF, Gajhede M. A flexible loop at the dimer interface is a part of the active site of the adjacent monomer of *Escherichia coli* orotate phosphoribosyltransferase. *Biochemistry*. 1996 Mar 26;35(12):3803-9.
- Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*. 1989 Apr 15;77(1):51-9.
- Hove-Jensen B, Bentsen AK, Harlow KW. Catalytic residues Lys197 and Arg199 of *Bacillus subtilis* phosphoribosyl diphosphate synthase. Alanine-scanning mutagenesis of the flexible catalytic loop. *FEBS J*. 2005 Jul;272(14):3631-9.
- Hove-Jensen B. Mutation in the phosphoribosylpyrophosphate synthetase gene (*prs*) that results in simultaneous requirements for purine and pyrimidine nucleosides, nicotinamide nucleotide, histidine, and tryptophan in *Escherichia coli*. *J Bacteriol*. 1988 Mar;170(3):1148-52.
- Howlett GJ, Minton AP, Rivas G. Analytical ultracentrifugation for the study of protein association and assembly. *Curr Opin Chem Biol*. 2006 Oct;10(5):430-6. Epub 2006 Aug 28.
- http://lowelab.ucsc.edu/GtRNAdb/Esch_coli_K12/Esch_coli_K12-by-isotype.html
- Huynh MT, Schilter D, Hammes-Schiffer S, Rauchfuss TB. Protonation of nickel-iron hydrogenase models proceeds after isomerization at nickel. *J Am Chem Soc*. 2014 Sep 3;136(35):12385-95. doi: 10.1021/ja505783z. Epub 2014 Aug 21.
- Iwasaki T. Iron-sulfur in aerobic and hyperthermoacidophilic archaea *Sulfolobus*. *Archaea*. 2010 Sep 19;2010. pii: 842639. doi: 10.1155/2010/842639.
- Javaid ZZ, el Kouni MH, Iltzsch MH. Pyrimidine nucleobase ligands of orotate phosphoribosyltransferase from *Toxoplasma gondii*. *Biochem Pharmacol*. 1999 Nov 1;58(9):1457-65.
- John RA. Pyridoxal phosphate-dependent enzymes. *Biochim Biophys Acta*. 1995 Apr 27;1248(2):81-96.
- Johnson DC, Dean DR, Smith AD, Johnson MK. Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem*. 2005;74:247-81.
- Jungbauer A, Hahn R. Ion-exchange chromatography. *Methods Enzymol*. 2009;463:349-71. doi: 10.1016/S0076-6879(09)63022-6.
- Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010; 66: 125-132. doi: 10.1107/S0907444909047337 PMID: 20124692
- Kadziola A, Jepsen CH, Johansson E, McGuire J, Larsen S, Hove-Jensen B. Novel class III phosphoribosyl diphosphate synthase: structure and properties of the tetrameric, phosphate-

Referencias

- activated, non-allosterically inhibited enzyme from *Methanocaldococcus jannaschii*. J Mol Biol. 2005 Dec 9;354(4):815-28. Epub 2005 Oct 18.
- Kessler D. Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes. FEMS Microbiol Rev. 2006 Nov;30(6):825-40.
- Kim S, Park S. Structural changes during cysteine desulfurase CsdA and sulfur acceptor CsdE interactions provide insight into de transpersulfuration. J Biol Chem. 2013 Sep 20;288(38):27172-80. doi: 10.1074/jbc.M113.480277. Epub 2013 Aug 2.
- Kimchi-Sarfaty C, Oh JM, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM. A 'silent' polymorphism in the MDR1 gene changes substrate specificity. Science. 2007 Jan 26;315(5811):525-8. Epub 2006 Dec 21.
- Kispal G, Csere P, Prohl C, Lill R. The mitochondrial proteins Atm1 and Nfs1p are essential for biogénesis of cytosolic Fe/S proteins. EMBO J. 1999 Jul 15;18(14):3981-9.
- Konarev P.V, Volkov V.V, Sokolova A.V, Koch M.H.J. and Svergun D.I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Cryst. (2003). 36, 1277-1282 doi:10.1107/S0021889803012779
- Krath BN, Hove-Jensen B. Implications of secondary structure prediction and amino acid sequence comparison of class I and class II phosphoribosyl diphosphate synthases on catalysis, regulation, and quaternary structure. Protein Sci. 2001 Nov;10(11):2317-24.
- Krissinel E. Macromolecular complexes in crystals and solutions. Acta Crystallogr D Biol Crystallogr. 2011; 67: 376-385. doi: 10.1107/S0907444911007232 PMID: 21460456
- Krungkrai SR, Aoki S, Palacpac NM, Sato D, Mitamura T, Krungkrai J, Horii T. Human malaria parasite orotate phosphoribosyltransferase: functional expression, characterization of kinetic reaction mechanism and inhibition profile. Mol Biochem Parasitol. 2004 Apr;134(2):245-55.
- Kulis-Horn RK, Persicke M, Kalinowski J. Histidine biosynthesis, its regulation and biotechnological application in *Corynebacterium glutamicum*. Microb Biotechnol. 2014 Jan;7(1):5-25. doi: 10.1111/1751-7915.12055. Epub 2013 Apr 25.
- Kumar S, Krishnamoorthy K, Mudeppa DG, Rathod PK. Structure of *Plasmodium falciparum* orotate phosphoribosyltransferase with autologous inhibitory protein-protein interactions. Acta Crystallogr F Struct Biol Commun. 2015 May;71(Pt 5):600-8. doi: 10.1107/S2053230X1500549X. Epub 2015 Apr 21.
- Lake MW, Wuebbens MM, Rajagopalan KV, Schindelin H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. Nature. 2001 Nov 15;414(6861):325-9.
- Land T, Rouault TA. Targeting of a human iron-sulfur cluster assembly enzyme, Nifs, to different subcellular compartments is regulated through alternative AUG utilization. Mol Cell. 1998 Dec;2(6):807-15.
- Laue TM. Sedimentation equilibrium as thermodynamic tool. Methods Enzymol. 1995;259:427-52.
- Laue TM, Shah BD, Ridgeway TM, Pelletier SL. Computer-aided interpretation of analytical

- sedimentation data for proteins. *Anal Ultracentrifugation Biochem Polym Sci*; 1992. pp. 90-125.
- Laue TM, Stafford WF 3rd. Modern applications of analytical ultracentrifugation. *Annu Rev Biophys Biomol Struct*. 1999;28:75-100.
- Lebowitz J, Lewis MS, Schuck P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci*. 2002 Sep;11(9):2067-79.
- Lehmann C, Begley TP, Ealick SE. Structure of the *Escherichia coli* ThiF-ThiS complex, a key component of the sulfur transfer system in thiamin biosynthesis. *Biochemistry*. 2006 Jan 10;45(1):11-9.
- Li J, Saxena S, Pain D, Dancis A. Adrenodoxin Reductase Homolog (Arh1p) of Yeast Mitochondria Required for Iron Homeostasis. *J Biol Chem*. 2001 Jan 12;276(2):1503-9.
- Li S, Lu Y, Peng B, Ding J. Crystal structure of human phosphoribosylpyrophosphate synthetase 1 reveals a novel allosteric site. *Biochem J*. 2007 Jan 1;401(1):39-47.
- Liu G, Li Z, Chiang Y, Acton T, Montelione GT, Murray D, Szyperski T. High-quality homology models derived from NMR and X-ray structures of *E. coli* proteins YgdK and SufE suggest that all members of the YgdK/SufE protein family are enhancers of cysteine desulfurases. *Protein Sci*. 2005 Jun;14(6):1597-608.
- Loiseau L, Ollagnier-de Choudens S, Lascoux D, Forest E, Fontecave M, Barras F. Analysis of the heteromeric CsdA-CsdE cysteine desulfurase assisting Fe-S clusters biogenesis in *Escherichia coli*. *J Biol Chem*. 2005 Jul 22;280(29):26760-9. Epub 2005 May 18.
- Loiseau L, Ollagnier-de-Choudens S, Nachin L, Fontecave M, Barras F. Biogenesis of Fe-S cluster by the bacterial Suf system. *J Biol Chem*. 2003 Oct 3;278(40):38352-9. Epub 2003 Jul 21.
- López-Esteva M, Ardá A, Savko M, Round A, Shepard WE, Bruix M, Coll M, Fernández FJ, Jiménez-Barbero J, Vega MC. The crystal structure and small-angle X-ray analysis of CsdL/TcdA reveal a new tRNA binding motif in the MoeB/E1 superfamily. *PLoS One*. 2015 Apr 21;10(4):e0118606. doi: 10.1371/journal.pone.0118606. eCollection 2015.
- López-Esteva M, Fernandez FJ, Vega MC. Needle in a haystack: targeting specific glucuronidases amid the human microbiome. *Curr. Enzyme Inhibition* 7(4), 218-228, DOI: 10.2174/157340811799860506
- Lorber B, Fischer F, Bailly M, Roy H, Kern D. Protein analysis by dynamic light scattering: methods and techniques for students. *Biochem Mol Biol Educ*. 2012 Nov-Dec;40(6):372-82. doi: 10.1002/bmb.20644. Epub 2012 Oct 10.
- Lorenz S, Cantor AJ, Rape M, Kuriyan J. Macromolecular juggling by ubiquitylation enzymes. *BMC Biol*. 2013 Jun 25;11:65. doi: 10.1186/1741-7007-11-65.
- Maarifi G, Maroui MA, Dutrieux J, Dianoux L, Nisole S, Chelbi-Alix MK. Small Ubiquitin-like Modifier Alters IFN Response. *J Immunol*. 2015 Sep 1;195(5):2312-24. doi: 10.4049/jimmunol.1500035. Epub 2015 Jul 29.

Referencias

- Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, Januszewski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, Helm M, Bujnicki JM, Grosjean H. MODOMICS: a database of RNA modification pathways--2013 update. *Nucleic Acids Res.* 2013 Jan;41(Database issue):D262-7. doi: 10.1093/nar/gks1007. Epub 2012 Oct 30.
- Matuszewski M, Sochacka E. Stability studies on the newly discovered cyclic form of tRNA N(6)-threonylcarbamoyladenine (ct(6)A). *Bioorg Med Chem Lett.* 2014 Jun 15;24(12):2703-6. doi: 10.1016/j.bmcl.2014.04.048. Epub 2014 Apr 21.
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40: 658–674. PMID: 19461840
- Mehta PK, Christen P. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol.* 2000;74:129-84.
- Mihara H and Esaki N. Bacterial cysteine desulfurases: their function and mechanisms. *Appl Microbiol Biotechnol.* 2002 Oct;60(1-2):12-23. Epub 2002 Sep 4.
- Mihara H, Kurihara T, Yoshimura T, Esaki N. Kinetic and mutational studies of three NifS homologs from *Escherichia coli*: mechanistic difference between L-cysteine desulfurase and L-selenocysteine lyase reactions. *J Biochem.* 2000 Apr;127(4):559-67.
- Miyauchi K, Kimura S, Suzuki T. A cyclic form of N6-threonylcarbamoyladenine as a widely distributed tRNA hypermodification. *Nat Chem Biol.* 2013 Feb;9(2):105-11. doi: 10.1038/nchembio.1137. Epub 2012 Dec 16.
- Moffatt BA, Ashihara H. Purine and pyrimidine synthesis and metabolism. *Arabidopsis Book.* 2002;1:e0018. doi: 10.1199/tab.0018. Epub 2002 Apr 4.
- Moulis JM, Davaise V, Golinelli MP, Meyer J, Quinkal I. The coordination sphere of iron-sulfur clusters: lesson from site-directed mutagenesis experiments. *JBIC Journal of Biological Inorganic Chemistry* February 1996, Volume 1, Issue 1, pp 2-14.
- Mueller EG. Trafficking in persulfides: delivering sulfur in biosynthetic pathways. *Nat Chem Biol.* 2006 Apr;2(4):185-94.
- Mühlenhoff U, Balk J, Richhardt N, Kaiser JT, Sipos K, Kispal G, Lill R. Nfs1p functional characterization of the eukaryotic cysteine desulfurase Nfs1p from *Saccharomyces cerevisiae*. *J Biol Chem.* 2004 Aug 27;279(35):36906-15. Epub 2004 Jun 25.
- Murphy FV 4th, Ramakrishnan V, Malkiewicz A, Agris PF. The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat Struct Mol Biol.* 2004 Dec;11(12):1186-91. Epub 2004 Nov 21.
- Ong HB, Sienkiewicz N, Wyllie S, Patterson S, Fairlamb AH. *Trypanosoma brucei* (UMP synthase null mutants) are avirulent in mice, but recover virulence upon prolonged culture *in vitro* while retaining pyrimidine auxotrophy. *Mol Microbiol.* 2013 Oct;90(2):443-55. doi: 10.1111/mmi.12376. Epub 2013 Sep 9.
- Outten FW, Djaman O, Storz G. A suf operon requirement for Fe-S cluster assembly during iron starvation in *Escherichia coli*. *Mol Microbiol.* 2004 May;52(3):861-72.

- Paiardini A, Contestabile R, D'Aguanno S, Pascarella S, Bossa F. Threonine aldolase and alanine racemase: novel examples of convergent evolution in the superfamily of vitamin B6-dependent enzymes. *Biochim Biophys Acta*. 2003 Apr 11;1647(1-2):214-9.
- Patching SG. Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. *Biochim Biophys Acta*. 2014 Jan;1838(1 Pt A):43-55. doi: 10.1016/j.bbamem.2013.04.028. Epub 2013 May 9.
- Percudani R, Peracchi A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep*. 2003 Sep;4(9):850-4.
- Petoukhov MV, Svergun DI. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys J*. 2005 Aug;89(2):1237-50. Epub 2005 May 27.
- Pflugrath JW. Practical macromolecular cryocrystallography. *Acta Crystallogr F Struct Biol Commun*. 2015 Jun;71(Pt 6):622-42. doi: 10.1107/S2053230X15008304. Epub 2015 May 27.
- Polyak SW, Abell AD, Wilce MC, Zhang L, Booker GW. Structure, function and selective inhibition of bacterial acetyl-CoA synthase. *Appl Microbiol Biotechnol*. 2012 Feb;93(3):983-92. doi: 10.1007/s00253-011-3796-z. Epub 2011 Dec 20.
- Prischi F, Pastore, A. Application of Nuclear Magnetic Resonance and Hybrid Methods to Structure Determination of Complex Systems. *Adv Exp Med Biol*. 2016;896:351-68. doi: 10.1007/978-3-319-27216-0_22.
- Putnam CD, Hammel M, Hura GL, Tainer JA. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys*. 2007 Aug;40(3):191-285.
- Raddatz G, Bisswanger H. Receptor site and stereospecificity of dihydrolipoamide dehydrogenase for R- and S-lipoamide: A molecular modeling study. *J Biotechnol*. 1997 Oct 17;58(2):89-100.
- Rathod PK, Khatri A, Hubbert T, Milhous WK. Selective activity of 5-fluoroorotic acid against *Plasmodium falciparum* *in vitro*. *Antimicrob Agents Chemother*. 1989 Jul;33(7):1090-4.
- Reddie KG, Carroll KS. Expanding the functional diversity of proteins through cysteine oxidation. *Curr Opin Chem Biol*. 2008 Dec;12(6):746-54. doi: 10.1016/j.cbpa.2008.07.028. Epub 2008 Sep 17.
- Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. Comparative genomics of the methionine metabolism in Gram-negative bacteria: a variety of regulatory systems. *Nucleic Acids Res*. 2004 Jun 23;32(11):3340-53. Print 2004.
- Rodriguez V, Chen Y, Elkahloun A, Dutra A, Pak E, Chandrasekharappa S. Chromosome 8 BAC array comparative genomic hybridization and expression analysis amplification and overexpression of TRMT12 in breast cancer. *Genes Chromosomes Cancer*. 2007 Jul;46(7):694-707.
- Round, A. The Use of Small-Angle Scattering for the Characterization of Multi Subunit Complexes. *Adv Exp Med Biol*. 2016;896:329-50. doi: 10.1007/978-3-319-27216-0_21.

Referencias

- Scapin G, Ozturk DH, Grubmeyer C, Sacchettini JC. The crystal structure of the orotate phosphoribosyltransferase complexed with orotate and alpha-D-5-phosphoribosyl-1-pyrophosphate. *Biochemistry*. 1995 Aug 29;34(34):10744-54.
- Scapin G. Molecular replacement then and now. *Acta Crystallogr D Biol Crystallogr*. 2013 Nov 1; 69 (Pt 11): 2266–2275. Published online 2013 Oct 18. doi: 10.1107/S0907444913011426
- Scherman MS, Kalbe-Bournonville L, Bush D, Xin Y, Deng L, McNeil M. Polyprenylphosphate-pentoses in mycobacteria are synthesized from 5-phosphoribose pyrophosphate. *J Biol Chem*. 1996 Nov 22;271(47):29652-8.
- Schröder M, Giermann N, Zrenner R. Functional analysis of the pyrimidine *de novo* synthesis pathway in solanaceous species. *Plant Physiol*. 2005 Aug;138(4):1926-38. Epub 2005 Jul 15.
- Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010
- Schulman BA, Harper JW. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat Rev Mol Cell Biol*. 2009 May;10(5):319-31. doi: 10.1038/nrm2673. Epub 2009 Apr 8.
- Schwarz G, Mendel RR, Ribbe MW. Molybdenum cofactors, enzymes and pathways. *Nature*. 2009 Aug 13;460(7257):839-47. doi: 10.1038/nature08302.
- Schweizer MP, Chheda GB, Baczynskyj L, Hall RH. Aminoacyl nucleosides. VII. N-(Purin-6-ylcarbamoyl)threonine. A new component of transfer ribonucleic acid. *Biochemistry*. 1969 Aug;8(8):3283-9.
- Sekowska A, Kung HF, Danchin A. Sulfur metabolism in *Escherichia coli* and related bacteria: facts and fiction. *J Mol Microbiol Biotechnol*. 2000 Apr;2(2):145-77.
- Sharma S, Sivalingam K, Neese F, Chan GK. Low-energy spectrum of iron-sulfur clusters directly from many-particle quantum mechanics. *Nat Chem*. 2014 Oct;6(10):927-33. doi: 10.1038/nchem.2041. Epub 2014 Aug 31.
- Shi R, Proteau A, Villarroja M, Moukadiri I, Zhang L, Trempe JF, Matte A, Armengod ME, Cygler M. Structural basis for Fe-S cluster assembly and tRNA thiolation mediated by IscS protein-protein interactions. *PLoS Biol*. 2010 Apr 13;8(4):e1000354. doi: 10.1371/journal.pbio.1000354.
- Shiba K, Niidome T, Katoh E, Xiang H, Han L, Mori T, Katayama Y. Polydispersity as a parameter for indicating the thermal stability of proteins by dynamic light scattering. *Anal Sci*. 2010;26(6):659-63.
- Shigi N. Biosynthesis and functions of sulfur modifications in tRNA. *Front Genet*. 2014 Apr 2;5:67. doi: 10.3389/fgene.2014.00067. eCollection 2014.
- Shimomura Y, Takahashi Y, Kakuta Y, Fukuyama K. Crystal structure of *Escherichia coli* YfhJ protein, a member of the ISC machinery involved in assembly of iron-sulfur clusters. *Proteins*. 2005 Aug 15;60(3):566-9.
- Singh H, Dai Y, Outten FW, Busenlehner LS. *Escherichia coli* SufE sulfur transfer protein modulates the SufS cysteine desulfurase through allosteric conformational dynamics. *J Biol Chem*. 2013 Dec 20;288(51):36189-200. doi: 10.1074/jbc.M113.525709. Epub 2013 Nov 6.

- Sprinzi M, Horn C, Brown M, Ioudovitch A, Steinberg S. Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.* 1998 Jan 1;26(1):148-53.
- Stellwagen E. Gel filtration. *Methods Enzymol.* 2009 ;463:373-85. doi: 10.1016/S0076-6879(09)63023-8.
- Sundaram M, Durant PC, Davis DR. Hypermodified nucleosides in the anticodon of tRNA^{Lys} stabilize a canonical U-turn structure. *Biochemistry.* 2000 Oct 17;39(41):12575-84.
- Svergun DI, Nierhaus KH. A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. *J Biol Chem.* 2000; 275: 14432-14439. PMID: 10799526
- Svergun DI. CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst.* 1995; 28: 768-773.
- Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 1999; 76: 2879-2886. PMID: 10354416
- Svergun, D. I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* 1992, 25, 495-503. doi: 10.1107/S0021889892001663.
- Switzer RL. Regulation and mechanism of phosphoribosylpyrophosphate synthetase. I. Purification and properties of the enzyme from *Salmonella typhimurium*. *J Biol Chem.* 1969 Jun 10;244(11):2854-63.
- Thiaville PC, Iwata-Reuyl D, de Crécy-Lagard V. Diversity of the biosynthesis pathway for threonylcarbamoyladenine (t(6)A), a universal modification of tRNA. *RNA Biol.* 2014;11(12):1529-39. doi: 10.4161/15476286.2014.992277.
- Tokumoto U, Kitamura S, Fukuyama K, Takahashi Y. Interchangeability and Distinct Properties of Bacterial Fe-S Cluster Assembly Systems: Functional Replacement of the *isc* and *suf* Operons in *Escherichia coli* with the *nifSU*-Like Operon from *Helicobacter pylori*. *J Biochem.* 2004 Aug;136(2):199-209.
- Torres AG, Batlle E, Ribas de Pouplana L. Role of tRNA modifications in human diseases. *Trends Mol Med.* 2014 Jun;20(6):306-14. doi: 10.1016/j.molmed.2014.01.008. Epub 2014 Feb 25.
- Trapani S, Navaza J. AMoRe: classical and modern. *Acta Crystallogr D Biol Crystallogr.* 2008 Jan;64(Pt 1):11-6. Epub 2007 Dec 5.
- Trotter V, Vinella D, Loiseau L, Ollagnier de Choudens S, Fontecave M, Barras F. The CsdA cysteine desulphurase promotes Fe/S biogenesis by recruiting Suf components and participates to a new sulphur transfer pathway by recruiting CsdL (ex-YgdL), a ubiquitin-modifying-like protein. *Mol Microbiol.* 2009 Dec;74(6):1527-42.
- Ucci JW, Cole JL. Global analysis of non-specific protein-nucleic interactions by sedimentation equilibrium. *Biophys Chem.* 2004; 108: 127-140. PMID: 15043926
- Van Hoewyk D, Abdel-Ghany SE, Cohu CM, Herbert SK, Kugrens P, Pilon M, Pilon-Smits EA. Chloroplast iron-sulfur cluster protein maturation requires the essential cysteine desulfurase CpnifS. *Proc Natl Acad Sci U S A.* 2007 Mar 27;104(13):5686-91. Epub 2007 Mar 19.

Referencias

Wang X, Ma C, Wang X, Xu P. Orotate phosphoribosyltransferase from *Corynebacterium ammoniagenes* lacking a conserved lysine. *J Bacteriol.* 2007 Dec;189(24):9030-6. Epub 2007 Oct 5.

Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, Matsui H, Atta M, Michiue H, Fontecave M, Yamagata K, Suzuki T, Tomizawa K. Deficit of tRNA(Lys) modifications by *Cdkal1* causes the development of type 2 diabetes in mice. *J Clin Invest.* 2011 Sep;121(9):3598-608. doi: 10.1172/JCI58056. Epub 2011 Aug 15.

White MN, Olszowy J, Switzer RL. Regulation and mechanism of phosphoribosylpyrophosphate synthetase: repression by end products. *J Bacteriol.* 1971 Oct;108(1):122-31.

Willemoës M, Hove-Jensen B, Larsen S. Steady state kinetic model for the binding of substrates and allosteric effectors to *Escherichia coli* phosphoribosyl-diphosphate synthase. *J Biol Chem.* 2000 Nov 10;275(45):35408-12.

Woolery AR, Luong P, Broberg CA, Orth K. AMPylation: something old is new again. *Front Microbiol.* 2010 Oct 19;1:113. doi: 10.3389/fmicb.2010.00113. eCollection 2010.

www.cpp4.ac.uk

www.expasy.org

www.genomics.agilent.com

www.hamptonresearch.com

www.ncbi.nih.gov/pubmed

www.phenix-online.org

www.rcsb.org

www.thermofisher.com

www.uniprot.org

www.xtal.iqfr.csic.es/Cristalografia/welcome.html

Yabe T, Morimoto K, Kikuchi S, Nishio K, Terashima I, Nakai M. The Arabidopsis chloroplastic NifU-like protein CnfU, which can act as an iron-sulfur cluster scaffold protein, is required for biogenesis of ferredoxin and photosystem I. *Plant Cell.* 2004 Apr;16(4):993-1007. Epub 2004 Mar 18.

Yanofsky C. RNA-based regulation of genes of tryptophan synthesis and degradation, in bacteria. *RNA.* 2007 Aug;13(8):1141-54. Epub 2007 Jun 29.

Yoo HM, Kang SH, Kim JY, Lee JE, Seong MW, Lee SW, Ka SH, Sou YS, Komatsu M, Tanaka K, Lee ST, Noh DY, Baek SH, Jeon YJ, Chung CH. Modification of ASC1 by UFM1 is crucial for ER α transactivation and breast cancer development. *Mol Cell.* 2014 Oct 23;56(2):261-74. doi: 10.1016/j.molcel.2014.08.007. Epub 2014 Sep 11.

Zheng M, Gu X, Zheng D, Yang Z, Li F, Zhao J, Xie Y, Ji C, Mao Y. UBE1DC1, an ubiquitin-activating enzyme, activates two different ubiquitin-like proteins. *J Cell Biochem.* 2008 Aug 15;104(6):2324-34. doi: 10.1002/jcb.21791.

Anexo

9. Anexo

Anexo I. Oligonucleótidos empleados en esta tesis.

	Oligonucleótido 5´	Oligonucleótido 3´	E. R.
CsdA	agg t ctcacatgaacgtttttaatcccgc	aaagg a tccttaatccaccaataattccagc	<i>BsaI/BamHI</i>
CsdAH120A	gcgtggcagaaacac gc ggccaacctcgctcc	gggacgaggttgg ccgc gtgttctgccacgc	-
CsdAK250A	ggctggg cg ggcg gc gatggttcacgaag	cttcgtgaaccat cg cgccgcccagcc	-
TcdA	aaac a tatgatgtctgtgtaattagtgatgca	aaaact cg agaccctgacgcgcc	<i>NdeI/XhoI</i>
tRNA ^{Lys}	aaaaa ga att cg ggctgtagctcagttgg	aaaact cg agtggtgggtcgtgcagg	<i>EcoRI/XhoI</i>
SufE	aaac a tatgatggctttattgccgataa	aaaact cg aggctaagtgcagcggtt	<i>NdeI/XhoI</i>
PRPP sintetasa	agg t ctcacatgcctgatatgaagcttttg	aaagg a tccttagtgttcgaacatggcaga	<i>BsaI/BamHI</i>
EcOPRT	aaac a tatgatgaaaccatatcagcgcca	aaaact cg agttaaacgcaaactcttcgc	<i>NdeI/XhoI</i>
T7 promotor	taatacgactcactataggg	-	-
T7 terminador	gctagttattgctcagcgg	-	-

Tabla 1. Oligonucleótidos empleados para el clonaje y mutagénesis para las distintas proteínas de esta tesis doctoral a excepción de los mutantes para CsdE. En todos los casos las secuencias se muestran desde el extremo 5´ hacia el 3´ de izquierda a derecha. También se señalan las enzimas de restricción (E.R.) empleadas para digerir los productos de PCR generados, así como su diana en la secuencia del oligonucleótido (negrita) o el codón mutado en el caso de los mutantes (negrita).

	Oligonucleótido 5´	Oligonucleótido 3´	E.R.
CsdE	aaac atat gatgacaaacccgcaattcg	aaaact cgaga acctgcttcgctgcag	<i>NdeI/XhoI</i>
R21E/N22E	ccgtaaccgcagaaacgtaga aagaa accttcgca ccgttgacgc	gcgtcaacggtgcgaagg ttctt ctaactgtt ctgcggttacgg	-
R86E	ctttggcgacagcgaaggg ga aattgtgcgcggc ctgctg	cagcaggccgcgcacaat ttcc cttcgctgtc gccaaag	-
R129E	cagcttagcgcctcag ag agccaggggttaaag	catttaaccctggct ctct gaggcgctaagctg	-
Y34E/R35E	gttgacgcaatgggaagataa aga aacagttg atcatgctggggaac	gtttcccagcatgatcaactg ttctt ctttatctt cccattgcgtcaac	-
K42E	gttgatcatgctgggg ga acagcttcggcattg	caatgccggaagctgt ttcccc cagcatgatcaac	-

Tabla 2. Oligonucleótidos empleados el clonaje de CsdE, así como de sus mutantes. Todos los oligonucleótidos mostrados y sus respectivas combinaciones dieron lugar a los mutantes de CsdE. En todos los casos las secuencias se muestran desde el extremo 5´ hacia el 3´ de izquierda a derecha. También se señalan las enzimas de restricción (E.R.) empleadas para digerir los productos de PCR generados al emplearlos, así como su diana en la secuencia del oligonucleótido (negrita) o el codón mutado en el caso de los mutantes (negrita). En el caso del mutante K42E/R86E, se empleó el mutante R86E como molde y los oligonucleótidos para K42E. En el caso del mutante K42E/R86E/R129E se empleó como molde K42E/R86E y los oligonucleótidos para R129E. En el caso del mutante Y34E/R35E/R129E se empleó como molde Y34E/R35E y los oligonucleótidos para R129E. Para el resto de mutantes se empleó como molde el gen codificante para el gen silvestre.

Publicaciones

RESEARCH ARTICLE

The Crystal Structure and Small-Angle X-Ray Analysis of CsdL/TcdA Reveal a New tRNA Binding Motif in the MoeB/E1 Superfamily

Miguel López-Esteva¹, Ana Ardá^{1,2}, Martin Savko³, Adam Round^{4,5}, William E. Shepard³, Marta Bruix⁶, Miquel Coll^{7,8}, Francisco J. Fernández¹, Jesús Jiménez-Barbero^{1,2}, M. Cristina Vega^{1*}



1 Chemical and Physical Biology, Center for Biological Research (CIB-CSIC), Madrid, Spain, **2** Structural Biology, CIC bioGUNE, Derio-Bizkaia, Spain, **3** Synchrotron SOLEIL, Gif-sur-Yvette, France, **4** European Molecular Biology Laboratory, Grenoble Outstation, Grenoble, France, **5** Unit for Virus Host-Cell Interactions, Univ. Grenoble Alpes-EMBL-CNRS, Martys, France, **6** Institute of Physical Chemistry "Rocasolano" (IFOR-CSIC), Madrid, Spain, **7** Institute for Research in Biomedicine (IRB Barcelona), Barcelona, Spain, **8** Institut de Biologia Molecular de Barcelona (IBMB-CSIC), Barcelona, Spain

* cvega@cib.csic.es

OPEN ACCESS

Citation: López-Esteva M, Ardá A, Savko M, Round A, Shepard WE, Bruix M, et al. (2015) The Crystal Structure and Small-Angle X-Ray Analysis of CsdL/TcdA Reveal a New tRNA Binding Motif in the MoeB/E1 Superfamily. PLoS ONE 10(4): e0118606. doi:10.1371/journal.pone.0118606

Academic Editor: Petri Kursula, Universitetet i Bergen, NORWAY

Received: November 26, 2014

Accepted: January 21, 2015

Published: April 21, 2015

Copyright: © 2015 López-Esteva et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The atomic coordinates and structure factors for the determined crystal structures are deposited in the Protein Data Bank (PDB) under accession numbers 4D79 (TcdA-ATP) and 4D7A (TcdA-AMP).

Funding: Ministerio de Economía y Competitividad (ES) (grants PET2008_0101, BIO2009-11184 and BFU2010-22260-C02-02 to MCV, BFU2008-02372/BMC, CONSOLIDER CSD 2006-23 and BFU2011-22588 to MC, CTQ2012-32035 to JJB), Generalitat de Catalunya (ES) (grant SGR2009-1309 to MC), the European Commission (Framework Programme 7

Abstract

Cyclic *N*⁶-threonylcarbamoyladenosine ('cyclic t⁶A', ct⁶A) is a non-thiolated hypermodification found in transfer RNAs (tRNAs) in bacteria, protists, fungi and plants. In bacteria and yeast cells ct⁶A has been shown to enhance translation fidelity and efficiency of ANN codons by improving the faithful discrimination of aminoacylated tRNAs by the ribosome. To further the understanding of ct⁶A biology we have determined the high-resolution crystal structures of CsdL/TcdA in complex with AMP and ATP, an E1-like activating enzyme from *Escherichia coli*, which catalyzes the ATP-dependent dehydration of t⁶A to form ct⁶A. CsdL/TcdA is a dimer whose structural integrity and dimer interface depend critically on strongly bound K⁺ and Na⁺ cations. By using biochemical assays and small-angle X-ray scattering we show that CsdL/TcdA can associate with tRNA with a 1:1 stoichiometry and with the proper position and orientation for the cyclization of t⁶A. Furthermore, we show by nuclear magnetic resonance that CsdL/TcdA engages in transient interactions with CsdA and CsdE, which, in the latter case, involve catalytically important residues. These short-lived interactions may underpin the precise channeling of sulfur atoms from cysteine to CsdL/TcdA as previously characterized. In summary, the combination of structural, biophysical and biochemical methods applied to CsdL/TcdA has afforded a more thorough understanding of how the structure of this E1-like enzyme has been fine tuned to accomplish ct⁶A synthesis on tRNAs while providing support for the notion that CsdA and CsdE are able to functionally interact with CsdL/TcdA.

(FP7) projects ComplexINC No. 279039 to MCV and SILVER-GA No. 260644 to MC). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Introduction

Transfer RNA (tRNA) molecules are targeted by more than 100 different enzymes that introduce a large number of diverse post-transcriptional modifications in tRNA nucleotides. This variety of modifications range from simple modifications at the base and/or at the 2'-hydroxyl of the ribose (including, among others, methylation, thiolation, deamination, and base isomerization) to more complex hypermodifications [1,2]. The functional roles of these modifications include the stabilization of the tertiary structure of tRNA, controlling gene expression, and modulating the interactions between tRNA and protein factors from e.g., the translational machinery [3]. Accordingly, tRNA modifications are essential for proper and efficient protein translation in all domains of life. Modifications in the anticodon loop of tRNA are crucial for decoding. A well-known example concerns the wobble modifications that occur at the first codon-anticodon position (U34), which participate in the regulation of cognate and near-cognate tRNA [3,4]. Hypermodified purine bases are frequently found 3' adjacent to the anticodon loop, at position 37 (A37) (Fig. 1A). Modifications at A37 have been associated with the stabilization of codon-anticodon interactions through base-stacking at the ribosomal decoding site.

N^6 -threonylcarbamoyladenine (t^6A ; Fig. 1B) is one of the 15 universally conserved essential modifications found in all three domains of life [2]. With the exception of bacterial initiator tRNA^{fMet}, t^6A occurs at position 37 (t^6A37) of the anti-codon stem loop (ASL) of all tRNAs which recognize ANN anticodons [5,6]. Further modifications of t^6A37 have been described; two tRNA species in *E. coli* have an N^6 -methyl derivative of t^6A (m^6t^6A) and a 2-methylthio derivative of t^6A (ms^2t^6A) is found in mammalian tRNA^{Lys} [7]. The breadth of functions fulfilled by t^6A37 in promoting the codon-anticodon interaction and accurate decoding by the ribosome spans aminoacylation of tRNAs [8], tRNA binding to the A-site codon [9,10], efficient

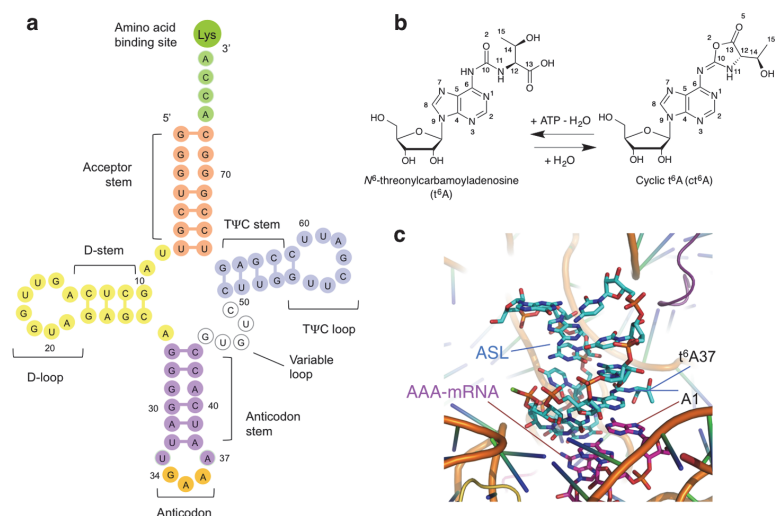


Fig 1. Cyclic N^6 -threonylcarbamoyladenine (cyclic t^6A , ct^6A) hypermodification in tRNA. a Schematic cloverleaf representation of tRNA^{Lys}UUU showing the location of the ct^6A modification at A37. **b** TcdA catalyzes the ATP-dependent cyclization of t^6A into ct^6A in tRNAs with ANN anticodons in bacteria, yeasts, fungi and plants, which occurs mainly at the A37 position of the anticodon stem loop (ASL) of the tRNA molecule. **(c)** Recognition of the AAA codon on mRNA by ASL of tRNA^{Lys}UUU involves interactions between t^6A37 and A1 from the AAA codon (PDB 1XMO).

doi:10.1371/journal.pone.0118606.g001

translocation [11], reading frame maintenance, preventing errors in AUG start codon selection and read-through of stop codons [12]. Steric hindrance from the bulky structure of t^6A impedes U33-A37 base pairing [13] so that the anticodon loop can assume its canonical U-turn structure [14]. The crystal structure of the codon-anticodon interaction at the A site of the 30S ribosomal subunit [15] has shown that the extended planar ring structure of t^6A37 (formed by intramolecular hydrogen bonding of the threonine moiety with the adenine ring) allows t^6A37 to stack with A38 in the anticodon loop and the first adenine base (A1) of the codon in the messenger RNA (mRNA) (Fig. 1C and S1 Fig.), thereby stabilizing the codon-anticodon interaction at the decoding site.

In bacteria, yeast, protists and plants the well-known t^6A modification has been recently shown to be a chemical hydrolysis artifact generated during the handling and preparation of tRNA [16]. Extraction of tRNA under extremely mild conditions has led to the identification of 'cyclic t^6A ' (ct^6A), a cyclized active ester with an oxazolone ring (Fig. 1B), as the bona fide hypermodification at position 37 of tRNA for ANN codons in those organisms [16]. The enzyme responsible for t^6A cyclization in *E. coli* was identified by comparative genomic approaches and analysis of LC/MS data sets of genomic deletion strains (ribonucleotide analysis) [17] as CsdL, subsequently renamed to TcdA, for tRNA threonylcarbamoyladenosine dehydratase A [16]. TcdA is an ubiquitin-activating E1-like protein [18] with detectable homology to the adenylyltransferases MoeB and ThiF. Previous studies have established that TcdA interacts with CsdE, a SufE-like sulfur acceptor [19], which in turn interacts with the cysteine desulfurase CsdA [20]. Indeed, ct^6A synthesis *in vivo* seems to depend on the presence of a functional CsdA-CsdE sulfur relay system since ct^6A abundance is drastically reduced in tRNA pools extracted from $\Delta csdE$ (2%) or $\Delta csdA$ (12%) cells [16]. Accordingly, CsdA alone or in concert with CsdE has been shown to support the enzymatic modification of TcdA by sulfur incorporation [18]. The relation between TcdA and CsdA-CsdE remains enigmatic, however, because ct^6A is a non-thiolated modification and the synthesis on a tRNA substrate of ct^6A starting from t^6A can be reconstituted *in vitro* with TcdA and ATP without CsdA or CsdE [16].

To start building a framework to investigate the molecular recognition processes underlying TcdA binding to ANN recognizing tRNAs and the structural and functional properties of TcdA responsible for the catalytic dehydration of t^6A into ct^6A , we set out to determine the structure of TcdA and its complex with tRNA^{Lys}(UUU). Toward this end, we have obtained high-resolution crystal structures of TcdA in complex with AMP and ATP, which show that the E1-like enzyme fold of TcdA exhibits a unique architecture that requires K^+ cations rather than Zn^{2+} for structural integrity (in contrast to most other E1-like proteins known) and a highly coordinated Na^+ cation at the dimer interface. Structural characterization by solution small-angle X-ray scattering (SAXS) of the TcdA complex with tRNA^{Lys}(UUU) reveals the position and relative orientation of the tRNA^{Lys}(UUU) substrate bound to TcdA. Finally, we have also probed into the structural and functional connection of TcdA with the CsdA-CsdE sulfur relay system using nuclear magnetic resonance (NMR) methods since CsdA and CsdE are essential components for ct^6A37 synthesis *in vivo* [16].

Materials and Methods

Recombinant TcdA production

E. coli TcdA (full length) was expressed in *E. coli* BL21 (DE3) cells harboring a C-terminal hexa-histidine tag. TcdA was purified using conventional nickel-chelating affinity chromatography (HisTrap, GE Healthcare) coupled to size-exclusion chromatography (Superdex 200 10/300GL, GE Healthcare) in 20 mM sodium/potassium phosphate, pH 7.4, 300 mM NaCl and 2 mM 2-mercaptoethanol.

Preparation of tRNA

The gene for *E. coli* tRNA^{Lys}(UUU) was cloned into the pET23a vector under the control of the strong T7 promoter, and expressed in *E. coli* BL21(DE3) cells grown for 3 h at 37°C after induction with 1 mM IPTG. The tRNA pool containing the over-expressed tRNA^{Lys}(UUU) was extracted from the cells by resuspending the cell pellet (from 200 ml culture) using acid phenol (pH 4.3), precipitated with 100% ethanol, and was further purified by size exclusion chromatography over a Superdex 75 HR 16/60 column (GE Healthcare) in phosphate buffer saline (PBS), 0.1 mM EDTA, pH 7.4. All buffers used for extraction and purification of tRNA^{Lys}(UUU) were twice autoclaved and the tRNA samples were kept at 4°C at all times to minimize the action of potential RNase contaminants. Agarose gel electrophoresis of the tRNA pool extracted from *E. coli* cells showed at least a 10-fold greater concentration upon induction when compared with non-induction controls, resulting in a tRNA pool highly enriched in recombinantly expressed tRNA^{Lys}(UUU).

Electrophoretic mobility shift assays with tRNA^{Lys}(UUU)

Gel retardation assays of TcdA complexed with over-expressed tRNA^{Lys}(UUU) or with commercial tRNA^{Lys}(UUU) (Sigma-Aldrich R1753), were performed by electrophoresing pre-incubated complexes and appropriate controls on native 8% polyacrylamide-0.5× Tris-borate-EDTA (TBE) gels at constant current (12 mA) for 1 h at 4°C. Gel retardation assays with the two sources of tRNA^{Lys}(UUU) gave identical results. Protein bands were stained with Coomassie Brilliant Blue and tRNA bands with Toluidine Blue.

Analytical ultracentrifugation

Sedimentation velocity (SV) and sedimentation equilibrium (SE) analytical ultracentrifugation (AUC) experiments were conducted in a Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge equipped with UV-Vis absorbance and Raleigh interference detection systems, using the 8-hole Beckman An-50Ti rotor at 20°C. TcdA, tRNA^{Lys}(UUU) and TcdA-tRNA^{Lys}(UUU) in 20 mM HEPES, 100 mM KCl and 50 mM NaCl, pH 7.4, were loaded (320 µl) into analytical ultracentrifugation cells. SV at 20 000 rpm was monitored by absorbance at 290 nm with scans made at 1 min intervals. Sedimentation coefficient distributions were calculated by least-squares boundary modeling of SV data using the continuous distribution $c(s)$ Lamm equation model as implemented in SEDFIT 14.1 [21]. Experimental s values were corrected to standard conditions (water, 20°C, and infinite dilution) in SEDNTERP to derive the corresponding standard s values ($s_{20,w}$). Short column (90 µl) SE experiments were carried out at speeds ranging from 8 000 to 12 000 rpm and monitored by absorbance at 280 nm, 290 nm and 295 nm. After the last equilibrium scans, a high-speed centrifugation run (48 000 rpm) was done to estimate the corresponding baseline offsets. Weight-average buoyant molecular weights of protein and nucleic acid species were determined by fitting a single species model to the experimental data using the HeteroAnalysis software [22,23] corrected for solvent composition and temperature with SEDNTERP [24].

TcdA crystallization, structure determination and refinement

Crystals of TcdA (full length, as a C-terminal hexa-histidine fusion) in complex with ATP or as obtained directly from the cell (AMP) were grown at 5 mg/ml TcdA in presence of 0.05 M potassium phosphate and 10% PEG1000 at basic pH at 20°C. The construct of TcdA and the crystallization conditions used in this study are different to those used in a recent crystallization report [25]. Two X-ray diffraction data sets were collected from crystals of the ATP complex,

one to a maximum resolution of 1.77 Å at a wavelength of 0.97948 Å, and a second data set to 2.35-Å resolution at a long wavelength (1.99976 Å) for sulfur single anomalous diffraction (SAD), at the BL13-XALOC beamline (ALBA, Barcelona, Spain). A data set for the AMP complex was collected to 1.89-Å resolution at the PROXIMA 2A beamline (Synchrotron SOLEIL,

Table 1. Crystallographic data processing and refinement statistics.

PDB code	TcdA-ATP 4D79	TcdA-AMP 4D7A
Data collection		
Wavelength (Å)	0.9795	0.9801
Resolution range (Å)	41.13–1.77 (1.83–1.77)	41.14–1.89 (1.96–1.89)
Space group	<i>P</i> 1 2 1 1	<i>P</i> 1 2 1 1
Unit cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	65.3, 96.7, 82.8	65.7, 97.2, 83.2
β (°), $\alpha = \gamma = 90^\circ$	90, 111.2, 90	90, 111.6, 90
Total reflections	312,887 (11,351)	146,386 (14,469)
Unique reflections	89,883 (5907)	87219 (78261)
Multiplicity	3.5 (1.9)	1.9 (2.0)
Completeness (%)	95.96 (63.90)	97.14 (92.68)
Mean $I/\sigma(I)$	18.97 (3.75)	18.99 (3.13)
Wilson B-factor	23.95	24.01
R-merge	0.08388 (0.663)	0.02326 (0.6354)
R-meas ^a	0.09797	0.327
CC1/2 ^b	0.995 (0.556)	0.878 (0.243)
CC ^c	0.999 (0.846)	0.967 (0.625)
Refinement		
R-work	0.1416 (0.2994)	0.1396 (0.1839)
R-free	0.1833 (0.3236)	0.1768 (0.2350)
# non-H atoms	8167	8064
# Protein atoms	7484	7411
# Ligand atoms	172	46
# Water	511	607
Protein residues	996	974
RMS(bonds) (Å)	0.011	0.007
RMS(angles) (°)	1.430	1.01
Ramachandran analysis		
Favored/Allowed/Outlier (%)	98.0/2.0/0.0	98.0/2.0/0.0
Clashscore	1.88	1.93
Average B-factor (Å ²)		
Protein	33.60	31.50
Ligands	60.80	49.70
Solvent	38.00	37.20

^aRmeas = $\sum_{hkl} (n/n-1)^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_i I_i(hkl)$; where *i* is the *i*th measurement of reflection (*hkl*) and $\langle I(hkl) \rangle$ is the average over symmetry related observations of a unique reflection (*hkl*).

^bCC1/2 is the Pearson correlation coefficient calculated between two random half data sets.

^cCC* is the CC of the full data set against the true intensities, estimated from $CC^* = [2 \text{ CC1/2} / (1 + \text{CC1/2})]^{1/2}$.

doi:10.1371/journal.pone.0118606.t001

Paris, France). All the data sets were integrated with XDS [26] and scaled with Aimless [27] from the CCP4 suite of programs [28] (Table 1).

The structure of TcdA in complex with ATP was solved at 1.77-Å resolution using MR-SAD by first placing residues 21–199 of MoeB (PDB 1JWA) [29] with PHASER [30] using the high-resolution data set, and then using the partial model and the anomalous signal of sulfur, phosphorus (from ATP) and potassium measured on the long wavelength data set to phase the full length structure. The complete structure of TcdA could be built into the minimally biased electron density maps calculated from MR-SAD phases; the only exception is a solvent exposed loop spanning residues 217–236 above the ATP binding pocket. The structure of the AMP complex was solved by MR using the ATP complex as search model. Both structures were built, refined and validated with *Coot* [31], phenix.refine [32] and MolProbity [33].

Accession codes

The atomic coordinates and structure factors for the determined crystal structures are deposited in the Protein Data Bank (PDB) under accession numbers 4D79 (TcdA-ATP) and 4D7A (TcdA-AMP).

Fluorescence spectroscopy

The dissociation constant for the TcdA-tRNA^{Lys}(UUU) binding interaction was obtained by following the quenching of the intrinsic fluorescence of tryptophan (2) and phenylalanine (8) residues in TcdA upon tRNA^{Lys}(UUU) addition. All fluorescence spectroscopy experiments (100 µl) were conducted in 20 mM sodium/potassium phosphate, pH 7.4, 300 mM NaCl, at 25°C using a Varioskan (Thermo Fisher) instrument, in black 96-well plates, setting the excitation and emission wavelengths to 280 nm and 340 nm, respectively. TcdA variants at 5 µM were challenged with increasing concentrations of tRNA (0.1–10 µM) in triplicate. Primary inner-filter effects due to tRNA absorption at the excitation wavelength were corrected for by standard methods (secondary inner filter effects for tRNA at 340 nm are negligible) and collisional induced quenching (as opposed to binding) was ruled out as the mechanism for the quenching of TcdA intrinsic fluorescence [34]. The reduction in fluorescence signal that accompanied tRNA addition was analyzed by non-linear regression methods in SigmaPlot v12, and the dissociation constant, K_D , and the Hill coefficient were calculated according to the following equation:

$$\frac{F_0 - F}{F - F_c} = \frac{L^n}{K_D + L^n}$$

where F_0 is the intrinsic fluorescence of TcdA in the absence of ligand, F_c is the minimal residual fluorescence of the TcdA-tRNA^{Lys}(UUU) complex and L is the tRNA^{Lys}(UUU) concentration.

NMR experiments

NMR experiments were acquired at 298 K on a Bruker Avance 600 MHz spectrometer equipped with a cryoprobe for 1D and 2D spectra, and on a Bruker Avance 800 MHz for the 3D. The spectra were processed with TopSpin 2.1 (Bruker), and 2D and 3D spectra were analyzed with the program CCPN Analysis [35]. 1D ¹H-¹³C and ¹H-¹⁵N HSQC spectra (5120 and 10240 scans respectively), 2D ¹H-¹⁵N HSQC spectra (3072 data points in t₂, 128 in t₁, 96 scans), and 3D spectra (2048 data points in t₃, 30 in t₂ (¹⁵N), and 128 in t₁ (¹³C), 32 scans) were acquired using standard Bruker pulse sequences. 1D heteronuclear spectra were processed with a line broadening of 10 Hz and 20 Hz (¹H-¹³C and ¹H-¹⁵N HSQC respectively). ¹H-¹⁵N

backbone resonance assignment for CsdE was carried out using HSQC, CBCANH [36] and CBCA(CO)NH [37] spectra and based on deposited data (BMRB code: 5630) [38].

Samples for the interaction between TcdA and CsdA were prepared as follows: a solution of 100 μ M of $^{13}\text{C}/^{15}\text{N}$ -TcdA (in binding buffer, consisting of 20 mM sodium/potassium phosphate, pH 7.4, 300 mM NaCl, 2 mM dithiothreitol) was divided in two parts: 100 μ l of one of them was diluted with 80 μ l of binding buffer and 20 μ l of D_2O , resulting in a 50 μ M solution of free TcdA. To the remaining 100 μ l, 80 μ l of a 1.3 mM solution of CsdA (unlabeled) and 20 μ l of D_2O were added, resulting in a 1:10 ratio solution TcdA:CsdA. 1D ^1H - ^{13}C and ^1H - ^{15}N HSQC spectra were acquired for both samples. 2D ^1H - ^{15}N HSQC spectra were acquired for samples 0.09 mM of $^{13}\text{C}/^{15}\text{N}$ -CsdE containing 123 μ l of binding buffer, and 0.09 mM $^{13}\text{C}/^{15}\text{N}$ -CsdE with 123 μ l of 0.6 mM TcdA (ratio 1:4).

Cross-linking experiments

The 7-atom, 8.0-Å short spacer arm cross-linking reagent bis(maleimido)ethane (BMOE) (Pierce 22322) generates non-cleavable cross-links between sulfhydryl groups that are in close proximity. BMOE was used to selectively trap transiently formed CsdE-TcdA complexes. CsdE and TcdA were mixed at a 1:1 molar ratio (80 μ M) in reaction buffer (20 mM sodium/potassium phosphate, 450 mM NaCl, 0.5 mM EDTA, pH 7.4), treated with 1 mM TCEP for 30 min to break pre-existing unspecific disulfide bridges, and BMOE was then added to 0.2 mM final concentration and incubated for 1 h at 25°C. Control reactions containing either CsdE or TcdA were set up and treated identically. The extent and specificity of the BMOE cross-linking reaction was followed by SDS-PAGE. The cross-linking reaction could be scaled up to obtain milligram amounts of CsdE-TcdA complex for biochemical and structural analysis (Supplementary Information).

SAXS measurements

SAXS experiments were performed at the BM29 BioSAXS beamline at the ESRF (Grenoble, France) [39]. SAXS data from purified TcdA were collected using a batch setup at three concentrations between 0.5–5 mg/ml, with ten successive time frames and 20 s exposures. For TcdA-tRNA^{Lys}(UUU), SAXS data (1 s per frame) were collected from two identical experiments using an online size-exclusion chromatography (SEC) setup [40] after injecting 100 μ l of 8-mg/ml complex on a Superdex 200 Increase column (GE Healthcare). Since the complex was prepared in presence of a molar excess of the tRNA^{Lys}(UUU) component, this step allowed to obtain scattering data from both the protein-tRNA complex and the excess tRNA^{Lys}(UUU) on a single experiment. All SAXS measurements were performed at 5°C in 20 mM sodium/potassium phosphate buffer, pH 7.4, 300 mM NaCl, 2 mM β -mercaptoethanol. Data were recorded using a 1 M PILATUS detector (DECTRIS) at a sample-to-detector distance of 2.7 m and a wavelength of 1.5 Å, covering the range of momentum transfer $0.020 < s < 0.5 \text{ Å}^{-1}$. Data from the batch setup (TcdA) or from the two equivalent peaks from the replicated SEC-SAXS measurements [TcdA-tRNA^{Lys}(UUU) and tRNA] were averaged, buffer subtracted and merged using the procedures outlined in Round et al. [40]. The radius of gyration (R_g) was evaluated using the Guinier approximation [41] and also from the entire scattering curve using Porod's law [42], and the pair-distance distribution function $P(r)$ was calculated using GNOM [43]. Guinier plots of all SAXS data support monodispersity of the analyzed samples (S2 Fig.). The measured scattering curves were compared with the theoretical scattering curves of the macromolecular models using CRY SOL [44]. *Ab initio* shape restoration was performed using 10–20 independent runs of DAMMIF [45] followed by DAMAVER [46] to create the final *ab initio* shape. ScÅtter was used to calculate the molecular weights of the macromolecules and for the

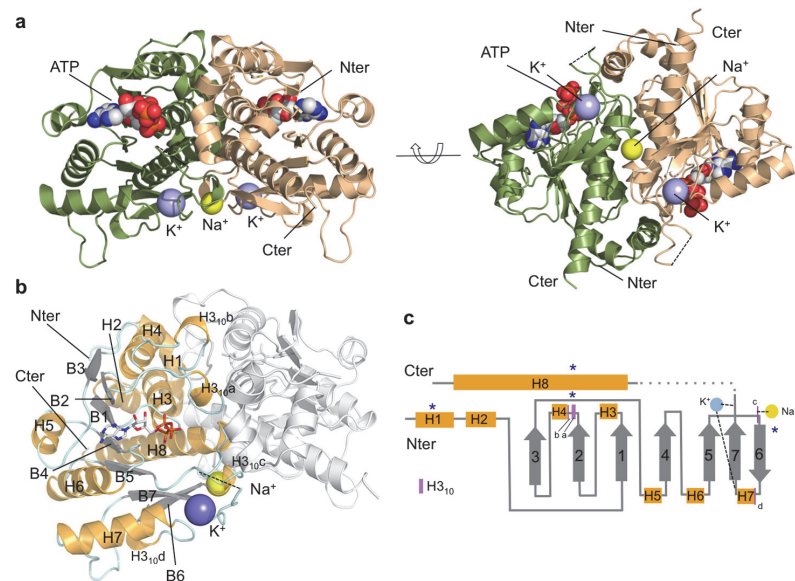


Fig 2. X-ray crystallographic structure of the binary complex TcdA at 1.77-Å resolution in complex with ATP. **a** Ribbon representation of the TcdA homodimer structure in two orientations related by a 90° rotation depicting each subunit in a different color (green or wheat). Two K⁺ and one Na⁺ cations are shown as violet and yellow spheres, respectively. The Na⁺ ion is located right at the homodimeric interface and its coordination sphere comprises amino acid residues from both monomers and water molecules. ATP is bound in a conserved surface pocket (shown as a space-filling model in CPK colors). **b** Description of secondary structural elements of TcdA monomer. Helices are shown in gold (α -helices are labeled H1-H8 and H_{310} helices H_{310a-d}), strands in white (B1-B7), and loops in light cyan. **c** Annotated schematic representation of TcdA topology. Helices and strands are depicted as in **b**, except for H_{310} helices which are indicated with purple lines and labeled a-d. K⁺ and Na⁺ ions are shown as blue and yellow circles. Dashed lines delimit regions that interact with K⁺, and structures involved in TcdA dimerization are marked with dark blue asterisks.

doi:10.1371/journal.pone.0118606.g002

calculation of the fitting parameter R_{SAS} , the small-angle scattering invariant V_C and the parameter Q_R [47]. The summary of SAXS statistics is given in [S1 Table](#).

Results and Discussion

Crystal structure of TcdA

The recent functional assignment of TcdA as the dehydratase in the synthesis of ct^6A was not predictable from the primary structure, since it has detectable sequence homology only to the ubiquitin-activating E1-enzymes MoeB (14%) and ThiF (11%). These enzymes are unknown to bind tRNA or catalyze post-translation modifications in tRNA substrates. The similarity to MoeB/ThiF is restricted to the N-terminal domain of TcdA (residues 30–157), which is classified into the ThiF family (Pfam database PF00899; E-value 6.7×10^{-41}) [48]. In contrast, the C-terminal end of TcdA lacks sufficient sequence similarity for functional prediction. To shed light onto the structural basis for the tRNA binding and ct^6A synthetic properties of TcdA-ATP, we determined the crystal structure of *E. coli* TcdA ([Table 1](#) and [Fig. 2](#)) loaded with ATP to 1.77 Å resolution (R/R_{free} values of 0.139/0.176) ([Fig. 3A](#)) and AMP to 1.89 Å resolution (R/R_{free} values of 0.141/0.183) ([Fig. 3B](#)). The asymmetric unit contained four TcdA chains arranged in two independent dimers, with a solvent content of 39%. Analysis of the

macromolecular interactions and potential complexes in the crystalline state with PISA [49] indicates that the TcdA dimers are the most likely assembly in solution. The structure of TcdA in complex with ATP was solved by molecular replacement-single wavelength anomalous diffraction (MR-SAD) using as search model residues 21–199 of MoeB (PDB 1JWA) [29] and the anomalous signal coming from protein sulfur atoms, ATP phosphorus atoms and K^+ cations as additional phasing information. The structure of TcdA-AMP was solved by MR using the fully refined model of the TcdA-ATP complex. Besides localized changes in the active site, the TcdA-ATP and TcdA-AMP structures are identical (RMSD 0.16 Å over 246 C α atoms).

The tertiary structure of monomeric TcdA consists of seven β -strands in a continuous sheet surrounded by eight α -helices (Fig. 2). All β -strands are parallel except for the sixth β -strand (β_6) at the end of the sheet (Fig. 2C). As in the MoeB/E1 superfamily, the N-terminal half of the sheet contains a variation of the Rossman fold whereby the $\beta\alpha\beta$ -topology is interrupted between β_2 and α_4 by the insertion of two consecutive 3_{10} helices. The first 3_{10} helix (H3_{10a}) contains three residues that are conserved in the MoeB/E1 superfamily, including a strictly conserved Asn69 that interacts with bound ATP/AMP via water-mediated hydrogen bonds. In agreement with its ATP hydrolytic activity, the loop between β_1 and α_3 contains the highly conserved Gly-X-Gly-Ala/Gly-Leu/Ile-Gly motif (where X denotes any residue, typically Val, Leu and Ile), akin to the P loop [50]. The C-terminal half of TcdA comprises three parallel β -strands (β_4 , β_5 and β_7) hydrogen bonded to the antiparallel β_6 , two α -helices connecting consecutive β -strands (α_6 and α_7), and a long α -helix (α_8) that runs perpendicular to the TcdA dimer symmetry axis and sticks out on either side (Fig. 2). Several of the C-terminal motifs including β_6 , α_7 and β_7 and the loops connecting them generate two unique metal binding sites, a K^+ binding site per TcdA monomer and the interfacial Na^+ binding site (Fig. 2).

In the TcdA-ATP complex (Table 1 and Fig. 3A), ATP is bound in a cleft over the C-terminal end of the central sheet. The residues in the P loop form the floor of the nucleotide triphosphate moiety, and the adenine ring is inserted into a hydrophobic cavity. Specific interactions anchor ATP at the active site through its triphosphate moiety and the ribose hydroxyl groups. A short hydrogen bond between the main-chain carbonyl oxygen from the P-loop Gly40 maintains the α -phosphate group firmly bound inside the binding pocket in a kinked conformation. The conserved Arg72 and Lys85 contact oxygens from the α - and β -phosphates, while Thr68 and the strictly conserved Asn69 in the H3_{10a} helix anchor the γ -phosphate through a water-mediated hydrogen bond. In the TcdA-AMP complex (Table 1 and Fig. 3B), the nucleotide-binding residues stabilize AMP in a manner analogous to ATP, with no conformational changes in the TcdA dimer. In the TcdA-ATP and TcdA-AMP complexes, the region between β_7 and the C-terminal α_8 spanning residues 214–236 is fully disordered in electron density maps. This surface disordered loop, in close proximity to the ATP binding site, corresponds to the disordered loop regions of MoeB [29] that are thought to play catalytic roles or engage in protein-protein interactions.

One K^+ binding pocket is found deeply buried in each TcdA monomer, as supported by anomalous Fourier maps of the K^+ ions contoured at 7 σ (Fig. 3C). The geometry of the coordination sphere of K^+ is distorted tetrahedral, and the ligand residues include main-chain carbonyl oxygens (Gln174, Ser206 and Glu208) and the side chain of Gln158, with an average metal-oxygen distance of 2.7 Å, and a water molecule at 2.6–3.1 Å. The K^+ binding site is created by structural elements in the C-terminal half of TcdA comprising the loop β_6 - α_7 , β_7 and loop- β_7 . This unique surface pocket is absent in other members of the MoeB/E1 superfamily and could have functional roles specific to tRNA binding and/or modification.

A second metal binding site specific for Na^+ is found at the dimer interface in an octahedral coordination (Fig. 3D). The Na^+ cation is axially liganded by the carbonyl oxygens of two symmetric Thr162 residues that lie in a short 3_{10} helix segment (H3_{10c}) from opposite monomers,

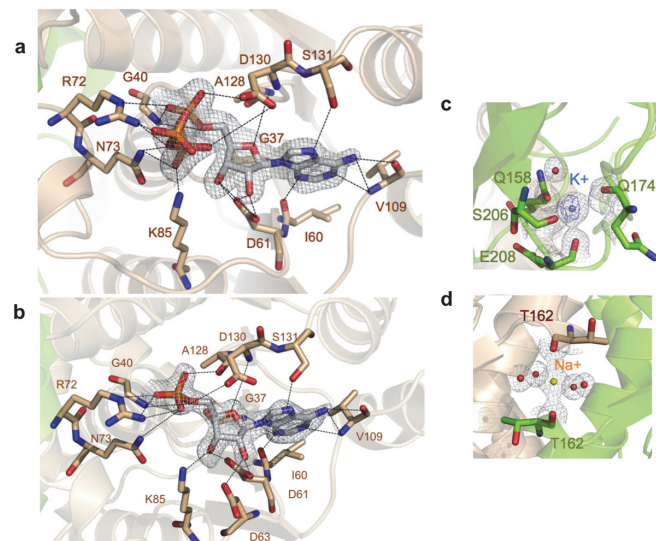


Fig 3. X-ray electron density maps of the ATP- and AMP-binding pocket and metal coordination spheres. **a** Detailed view of the ATP-binding site of TcdA. Residues that interact with ATP are labeled and shown as sticks and atom colors (C atoms are in subunit colors, as in Fig. 1). $2mF_o - DF_c$ electron density map is depicted around the ATP substrate at 1.5σ contour level. **b** Detailed view of AMP bound in the active site. Representation and electron density map as in (a). **c, d** Metal coordination spheres of K^+ (c) and Na^+ (d) cations. Interacting residues are labeled and shown as sticks in atom colors (C atoms are in subunit colors), and coordinating water molecules are shown as red spheres. Metals and their coordination spheres are shown in $2mF_o - DF_c$ electron density (grey). The anomalous Fourier map calculated from a long wavelength (1.89 Å) dataset at 2.4-Å resolution is shown for K^+ (c) in purple; in contrast, Na^+ has no anomalous signal at that wavelength, enabling the accurate discrimination between K^+ and Na^+ cations in the structure.

doi:10.1371/journal.pone.0118606.g003

with four water molecules occupying the equatorial positions. The identity of this sodium ion was unambiguous owing to the conspicuous coordination sphere of sodium, with average coordination distances of 2.35 Å between the metal and the carbonyl oxygens, and 2.50 Å with the water oxygens, and the total absence of anomalous signal at 1.999 Å, which ruled out a K^+ .

Dimeric structure of TcdA

The overall architecture of TcdA consists in a symmetric dimer, consistent with the observed elution volume of TcdA in analytical gel filtration (not shown) and solution SAXS scattering data (S1 and S2 Tables). The TcdA dimer is maintained by Na^+ , which mediates monomer-monomer contacts through the dimer interface, and by K^+ ions, which play a structural role in shaping the monomer fold. The dimer interface buries approximately 2250 Å² solvent accessible surface area (ASA), or 20% of the total ASA per monomer, engaging >25% of TcdA residues along an elongated surface patch (Fig. 4A). According to PISA [49], the TcdA dimer is chemically stable as its free Gibbs energy of dissociation is 21.5 kcal/mol. This surface patch comprises four helices ($\alpha 1$, $\alpha 3$, $\alpha 4$ and the long C-terminal helix $\alpha 8$) from each monomer, the 3_{10} helical regions H3_{10a} and H3_{10c} that shape the nucleotide-binding site and the Na^+ binding site, respectively, and the loops $\beta 5$ - $\beta 6$ and $\beta 7$ - $\alpha 8$. In addition to many hydrophobic interactions, several strong hydrogen bonds stabilize the dimer structure, engaging interfacial arginine residues from helices $\alpha 1$ (Arg12), $\alpha 3$ (Arg51), H_{10b} (Arg78) and $\alpha 4$ (Arg92) (Fig. 4B,C). Two symmetric interactions further contribute to the dimer stability, the antiparallel interaction

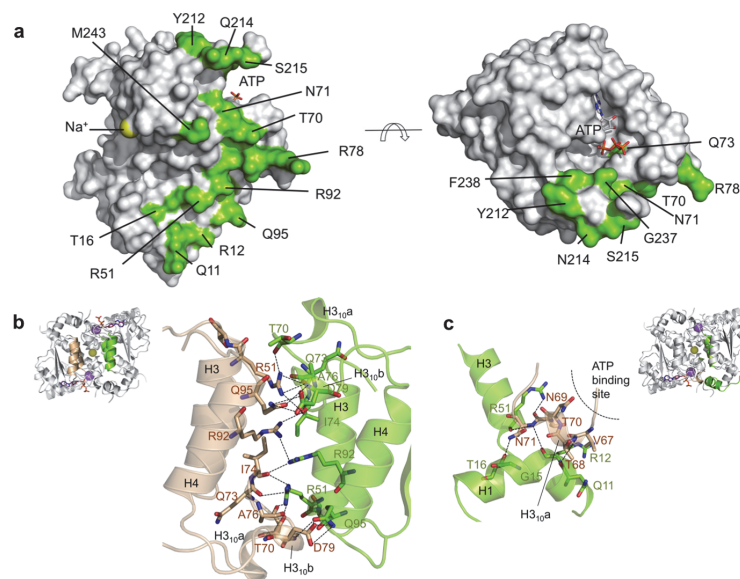


Fig 4. TcdA dimerization interface. Molecular surface representation of a TcdA dimer in two orientations (a) related by a 90° rotation around a horizontal axis. The location of the ATP binding pocket and the interfacial Na⁺ ion are labeled. Amino acid residues that contribute to the extended, flat dimer interface are mapped onto the molecular surface in green and labeled. **b** Close-up of the symmetric helical bundle (H3-H4). Key hydrogen bonding and charge interactions are shown as dashed lines. **c** Helix H1 participates in the dimer interface through hydrogen bonding and van der Waals interactions with H3 from the same chain and H3_{10a} from the opposite monomer.

doi:10.1371/journal.pone.0118606.g004

between helices $\alpha 8$ and the Na⁺-bridged interaction held by Thr162 main-chain oxygens (Fig. 3D).

Comparison with MoeB/E1 superfamily

There is significant structural similarity between the 180-amino-acid N-terminal half of TcdA and the catalytic domain of MoeB/ThiF, which can be superimposed with RMSD of 1.57 Å (181 C α atoms) and 1.42 Å (177 C α atoms), respectively (Fig. 5). The MoeB/E1 superfamily enzymes *E. coli* MccB and human UBA5, which share the same fold structure than MoeB/ThiF, also display a considerable degree of structural similarity to TcdA, with RMSD of 1.74 Å (173 C α atoms) and 1.85 Å (146 C α atoms), respectively (Fig. 5). The nucleotide-binding pocket and a significant part of the helices that in TcdA participate in the dimer interface are among the most conserved structures between TcdA and MoeB/ThiF. In particular, the P loop and the residues that contact ATP/AMP are nearly strictly conserved across all structural homologs.

In contrast to the conservation displayed by the nucleotide-binding site and adjacent regions, the C-terminal half of TcdA adopts a unique structure that is distinct from those present in MoeB/ThiF. Firstly, the C-terminal extension of TcdA is stabilized by a structural K⁺ cation and not by Zn²⁺, which is consistent with the absence in TcdA of the Cys4 motif responsible for Zn²⁺ binding in MoeB/ThiF. Of the four cysteine residues used by MoeB/ThiF to chelate Zn²⁺, TcdA has retained only the last cysteine (Cys220), while the first two cysteine residues fall in a sequence stretch absent in TcdA and the third cysteine is a threonine in TcdA (Thr218). Cys220 is situated in the beginning of the disordered loop and therefore could have a

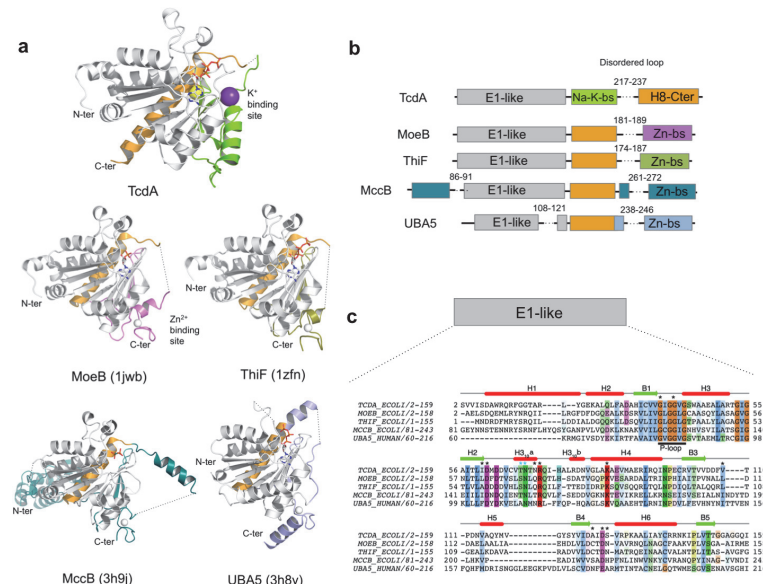


Fig 5. Structural comparison of TcdA with the homologous E1-like activating enzymes. **a** Ribbon representation of TcdA structural homologs found with PDBFold [51] with structural similarity Q-scores higher than 0.28 (the highest Q-scores are 0.49 with MoeB and 0.46 with ThiF). All structures were superimposed onto TcdA and are shown in similar orientations for comparison: MoeB (PDB 1JWB), ThiF (PDB 1ZFN), MccB (PDB 3H9J) and UBA5 (PDB 3H8V). E1-like core domain is shown in grey; the long α -helix that is structurally equivalent to TcdA $\alpha 8$ is in orange; metal binding sites are in protein-specific colors; K^+ ions are in purple, Na^+ ions in yellow, and Zn^{2+} in grey. **b** Schematic representation of the domain architecture of the E1-like enzymes superimposed in (a), with equivalent color coding; bs, binding site; the sequence position corresponding with disordered loops is indicated by dashed lines and marked with start and end residues. The C-terminal $\alpha 8$ helix in TcdA is structurally equivalent to the α -helix immediately following the E1-like domain in the other enzymes. **c** Structure-guided multiple sequence alignment of the E1-like domain with overlaid secondary structure from TcdA. Functionally important residues and motifs are annotated as follows: bold underline, P-loop residues; black asterisks, residues in direct contact with ATP; and blue asterisks, residues from helix H3_{10a} that make water-mediated contacts with the nucleotide. Conserved residues are shown in shaded colors.

doi:10.1371/journal.pone.0118606.g005

functional role during catalysis. Cys234 in TcdA is also conserved in MoeB/ThiF, and it is found in the disordered loop. In contrast, an additional Cys66 without an equivalent residue in MoeB/ThiF is located near the active site (S3 Fig). Since functional thiol transfer from CsdA-CsdE appears to be essential for TcdA function *in vivo* [16], and thiolation of TcdA has been shown to occur *in vitro* [18], it seems likely that Cys66, Cys220 or Cys234 may be susceptible to be the recipient of the activated sulfur atom.

As a result of the structural differences in the C terminus of TcdA and MoeB/ThiF, the solvent-exposed surface motifs that allow the MoeB/MoeD and ThiF/ThiS complexes to form are blocked in TcdA by an extended segment containing residues 170–189, which is part of the K^+ binding site. This motif is the only structural element that has no counterpart in MoeB and ThiF, raising the possibility that it is involved in TcdA specific functional roles absent in other E1-like enzymes. In addition, structure superposition of TcdA with MoeB, ThiF, MccB and UBA5 (Fig 5 and S4 Fig) shows a difference in domain organization consisting in a swap between the long helix (orange) and the metal binding sites. In TcdA, in contrast to all other

structural homologs, the metal binding site occupies a central position between the E1-like domain and the C-terminal $\alpha 8$.

TcdA can establish transient interactions with CsdE and CsdA

Previous studies have established that TcdA can be persulfurated by the cysteine desulfurase CsdA, either directly or indirectly through the SufE-like sulfur acceptor CsdE [18], and that the *csdA* and *csdE* genes are essential to support the function of TcdA in *ctA* biosynthesis *in vivo* but not in an *in vitro* reconstituted system [16]. To shed light into the proposed interaction between TcdA and CsdA-CsdE we set out to characterize their complex. We were not able to identify a stable CsdA-TcdA or CsdE-TcdA complex by gel filtration or pull-down experiments (data not shown) even though a CsdE-TcdA complex had been previously isolated using over-expressed TAP-tagged CsdE [18]. To investigate whether a short-lived complex of TcdA with either CsdE or CsdA could form *in vitro*, we resorted to 1D and 2D HSQC NMR techniques that can detect fast exchanging transient protein interactions with high sensitivity.

To probe the hypothetical interaction between CsdA and TcdA, which would result in a 148-kDa heterotetramer (assuming one TcdA monomer bound per subunit of CsdA homodimer), we acquired 1D ^1H - ^{13}C and ^1H - ^{15}N HSQC spectra of $50\ \mu\text{M}$ ^{13}C , ^{15}N -TcdA (20% D_2O) upon addition of unlabeled CsdA at a 10-fold excess molar ratio. Comparison of the resultant spectrum with a blank experiment recorded for only free isotope-labeled TcdA permitted to detect a decrease in the signal intensity of the TcdA resonances in the presence of ten-fold CsdA (data not shown). This decrease in intensity suggests the existence of a very weak interaction between CsdA and TcdA, consistent with published data reporting a catalytic interaction for sulfur transfer between TcdA and CsdA and the absence of a stable CsdA-TcdA complex by a yeast two-hybrid analysis [18].

In contrast to the CsdA-TcdA complex, the smaller size of CsdE and the availability of its solution structure by NMR (bmr5630, PDB code 1NI7) [38] afforded the opportunity to further characterize the region of its putative interaction with TcdA. In this case, regular ^1H - ^{15}N 2D-HSQC experiments (Fig. 6A) using ^{13}C , ^{15}N -doubly labeled CsdE and native TcdA were performed. In these experiments, the ^1H - ^{15}N HSQC spectra of free ^{13}C , ^{15}N -CsdE ($90\ \mu\text{M}$) and of a similar sample with 4-fold molar excess of unlabeled TcdA were acquired. The comparison of these spectra showed that several CsdE crosspeaks suffered chemical shift perturbations in the presence of TcdA, while most of them remained unperturbed. These evidences demonstrate the existence of a transient physical interaction between CsdE and TcdA. Given the observed chemical shift perturbations, the molecular recognition process is in fast exchange in the NMR chemical shift time scale, with an estimated K_D in the mM to high μM range. Under the low excess ratio conditions (1:4) between the components, only a small fraction of CsdE is probably bound in the CsdE-TcdA complex. Thus, the existence of a fast exchange process permits explaining the absence of complete signal decay. The assignment of the crosspeaks that were perturbed by TcdA allowed mapping the CsdE surface in contact with TcdA (Fig. 6B,C). In fact, the most perturbed region was located at residues 70–81, corresponding to the antiparallel hairpin that follows Cys61, the persulfurated Cys in CsdE, allowing its identification as the most crucial for TcdA interaction (Fig. 6C). This region has been characterized as being rather dynamic, as shown by both the CsdE and CsdA-CsdE crystal structures [38,52]. Very probably, the elongated structure of CsdE can facilitate further intermolecular interactions with the surface-exposed cysteine residues in TcdA.

To seek further support for a transient CsdE-TcdA complex, we performed cross-linking experiments with bis(maleimido)ethane (BMOE), a compound that cross-links sulfhydryl groups that are in close proximity. The short 7-atom spacer arm of BMOE spans an 8.0-Å distance

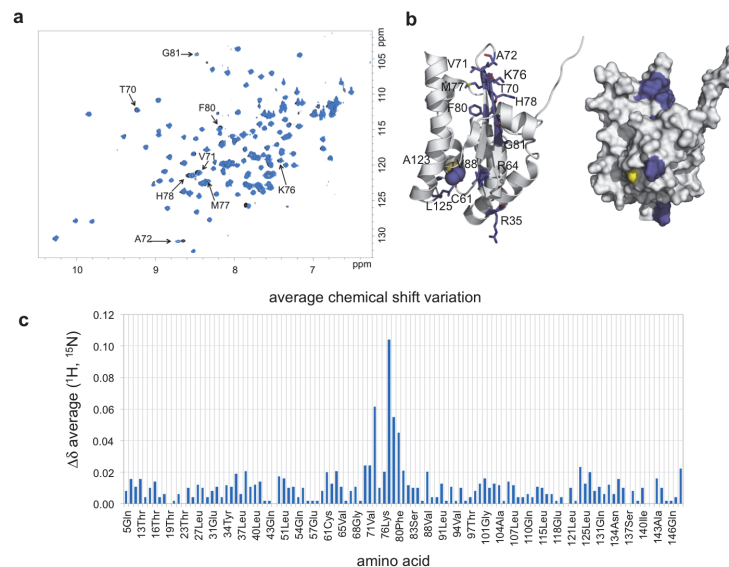


Fig 6. TcdA interacts transiently with the sulfur-acceptor SufE-like protein CsdE. **a** NMR ^1H - ^{15}N NOESY spectrum of doubly labeled ^{13}C , ^{15}N -CsdE in presence (blue) or absence (black) of unlabeled TcdA. Significant shifts in the position of specific NOESY resonance peaks for CsdE residues are labeled. **b** The position of the affected amino acid residues is mapped onto the NMR structure of CsdE (PDB 1NI7) in ribbon (left) and surface (right) representations. **c** Average chemical shift ($\Delta\delta$) of ^{13}C , ^{15}N -CsdE upon binding to TcdA as measured from the NOESY spectrum in **a** plotted against the CsdE amino acid sequence.

doi:10.1371/journal.pone.0118606.g006

between the two cross-linked thiol groups, therefore enabling the capture of specific protein complexes. Our cross-linking experiments demonstrated the selective cross-linking of CsdE and TcdA in SDS-PAGE, with no cross-linked species being detected for CsdE or TcdA alone. The cross-linked CsdE-TcdA complex was stable against incubation with DTT and could be separated from excess components by gel filtration. The analysis of the cross-linked complex by SAXS revealed conformational flexibility that could be interpreted in terms of two independent CsdE molecules covalently attached to TcdA in pseudo-symmetrical positions (S1 Text and S5 Fig.).

Highly specific cross-linking of CsdE and TcdA with BMOE lends further support to the notion that CsdE and TcdA might engage in a complex. These weak, fast exchanging interactions may be functionally important for TcdA function provided that TcdA post-translational modification by persulfuration are required for proper function in ct^6A biosynthesis.

TcdA forms a complex with tRNA^{Lys}

The proposed role for TcdA in the transformation of $\text{t}^6\text{A}37$ into its cyclic analogue, $\text{ct}^6\text{A}37$, at position 37 of tRNA^{ANN} , requires that TcdA engage in a productive binding interaction with its cognate tRNAs. Since no previous evidence for such a complex was available, we conducted electrophoretic mobility shift assays (EMSA) with purified TcdA and $\text{tRNA}^{\text{Lys}}(\text{UUU})$ in the presence of the co-substrate Mg^{2+} -ATP (Fig. 7C). These assays revealed the formation of a TcdA- $\text{tRNA}^{\text{Lys}}(\text{UUU})$ complex that migrated as a distinct band when compared with TcdA alone (which did not enter the gel owing to its basic pI of 8.8) or isolated $\text{tRNA}^{\text{Lys}}(\text{UUU})$,

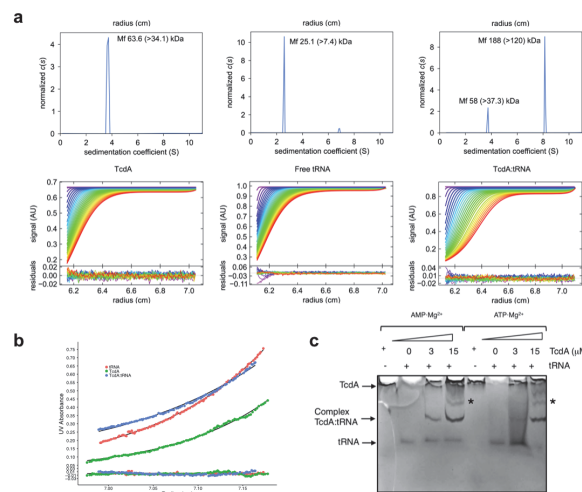


Fig 7. TcdA binding to tRNA. **a** SV-AUC experiments demonstrate that TcdA contains a tRNA^{Lys} binding domain. Sedimentation coefficient (S) continuous distributions, $c(S)$, plotted for TcdA, tRNA^{Lys}, and the TcdA-tRNA^{Lys} complex. Representative sedimentation scans and residuals after fitting are shown. For each $c(S)$ distribution, the estimated molecular mass and the smallest mass compatible with the experimental results are indicated. **b** Absorbance at 280 nm plotted against the AUC column length for SE-AUC experiments for TcdA, tRNA^{Lys}, and TcdA-tRNA^{Lys}. Residuals after fitting are also shown. **c** Electrophoretic mobility shift assay (EMSA) provides evidence for the physical binding of TcdA to tRNA^{Lys}. Samples were run over a 8% TBE-PAGE gel at 4°C for 1 h to fully resolve tRNA^{Lys} from TcdA-tRNA^{Lys} bands. TcdA does not enter the gel when tRNA^{Lys} is not added, most likely due to its intrinsically basic isoelectric point, pI (theoretical $pI = 8.8$). tRNA^{Lys} alone migrates as the fastest migrating band. When 5 M tRNA^{Lys} is added to 3 μ M or 15 μ M TcdA, protein-tRNA^{Lys} complexes develop that migrate behind free tRNA^{Lys}. An arrow indicates the position of the retardation band containing the most abundant TcdA-tRNA^{Lys} complex (2:2 stoichiometry), and an asterisk marks the position of a slower migrating TcdA-tRNA^{Lys} complex (most likely, 2:1 stoichiometry).

doi:10.1371/journal.pone.0118606.g007

which migrated at a lower apparent molecular size. At the highest concentration of TcdA (15 μ M), additional higher molecular weight bands become discernable, which may represent higher-order complexes. Although the physiologic quaternary structure of the TcdA-tRNA^{Lys}(UUU) complex remains unknown, the dimeric structure of TcdA suggests that the fastest migrating complex retardation band may contain two tRNA^{Lys}(UUU) molecules in a 2:2 stoichiometric complex (Fig 7C, arrow) while the next slower migrating band might contain one tRNA^{Lys} molecule in a 2:1 substoichiometric complex with TcdA (Fig 7C, *). The K_D for the TcdA-tRNA^{Lys}(UUU) complex was determined as 3 μ M by using the fact that the tryptophan fluorescence of TcdA is quenched by tRNA.

We further characterized the size and stability of the TcdA-tRNA^{Lys}(UUU) complex by performing sedimentation velocity (SV) and sedimentation equilibrium (SE) analytical ultracentrifugation (AUC) experiments using complexes assembled in presence of a molar excess of the tRNA^{Lys}(UUU) component. SV-AUC profiles for the complex was consistent with the formation of a high-molecular weight protein-tRNA complex that sedimented with a sedimentation coefficient of 8.2 S, compared with the isolated components (3.8 S for TcdA and 2.4 S for the tRNA) and with remaining free TcdA in the complex experiment (3.8 S) (Fig 7A). Using the same experimental conditions for the previous experiments, we then ran SE-AUC for the TcdA-tRNA^{Lys}(UUU) complex to measure the buoyant mass of the complex and assess the co-existence of excess individual components (Fig 7B). The calculated molecular masses for the sedimenting species are within experimental error to the theoretical masses for all components:

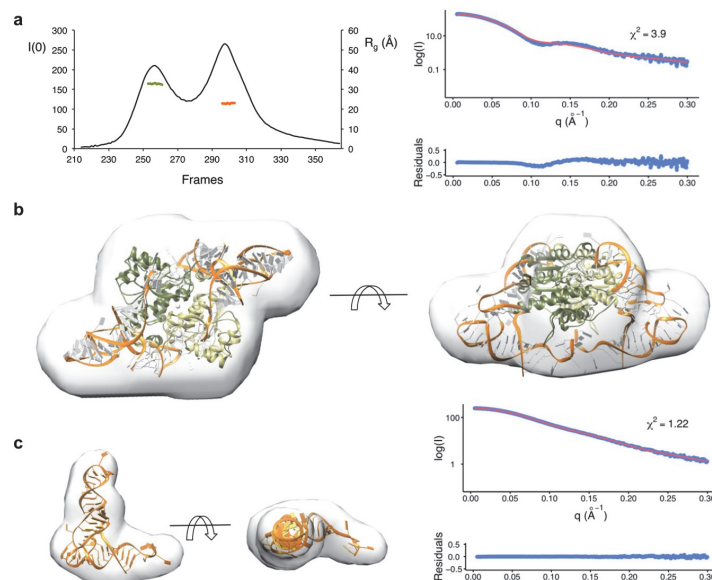


Fig 8. Solution structure of TcdA-tRNA^{Lys}(UUU). **a** SAXS data for TcdA-tRNA^{Lys} was measured using an online HPLC setup to separate complex from excess free tRNA. The graph shows a plot of the SAXS intensity at zero angle, $I(0)$ (left axis, curve represented as a solid black line), and of the radius of gyration, R_g (on the right axis), versus data-collection frames. Frames 253–263 (green line) were merged and used for shape restoration of the TcdA-tRNA^{Lys}(UUU), and frames 304–310 were used for the control reconstruction of the tRNA^{Lys}(UUU) shape. **b** Best model calculated for the TcdA-tRNA^{Lys}(UUU) complex overlaid by the ab initio SAXS shape calculated with DAMMIF. The crystal structure of TcdA-ATP is represented in green cartoon and the tRNA is depicted with its main chain as a gold ribbon and the bases as ladders. The fit (red line) to the experimental SAXS data (blue points), calculated with CRY SOL, χ^2 and residuals are shown. **c** Like in **(b)**, for free tRNA^{Lys}(UUU). In this case, the model is a rigid-body fit of the tRNA structure into the ab initio SAXS envelope.

doi:10.1371/journal.pone.0118606.g008

63,531 ± 840 Da (TcdA dimer, 57.9 kDa), 26,066 ± 156 Da (tRNA, 23.4 kDa), and 114,297 ± 824 Da [TcdA-tRNA^{Lys}(UUU), 104.8 kDa] (S2 Table).

SAXS reconstruction of TcdA-tRNA^{Lys} complex

In the absence of a crystallographic structure for the TcdA-tRNA^{Lys}(UUU) complex, and in view of the dynamic nature of the complex as inferred from gel filtration and AUC experiments, we resorted to shape restoration by SAXS to elucidate the three-dimensional structure of the TcdA-tRNA^{Lys}(UUU) complex. To collect X-ray scattering data from a monodisperse, homogenous solution of the complex, and circumvent the dynamic dissociation of the complex into its constituents, we subjected pre-incubated complex formed from purified components to on-line HPLC coupled with SAXS. The setup allowed us to collect useful solution scattering data from both the complex and from the excess tRNA^{Lys}(UUU) (Fig. 8A).

Seven to fifteen successive frames across the non-overlapping peaks from a duplicated experiment containing TcdA-tRNA^{Lys}(UUU) or free tRNA^{Lys}(UUU) were processed and merged together on the basis of a coherent experimental R_g (R_g complex: 33.4 Å [Guinier points 24–53; s R_g 0.471–0.930; Fidelity 0.706]; R_g free tRNA: 23.0 Å [Guinier points 5–72; s R_g 0.139–0.892; Fidelity 0.979]). SAXS invariants were calculated for these species (reported in S1 Table).

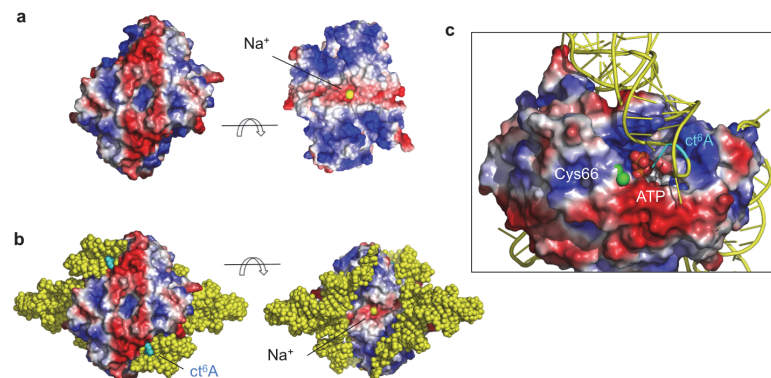


Fig 9. TcdA-tRNA interface. **a** Electrostatic potential molecular surface calculated with APBS (Adaptive Poisson-Boltzmann Solver) [55] and rendered with PyMOL (www.pymol.org) [56]. Two views are shown that are related by a 90° rotation around a horizontal axis. The interfacial Na⁺ cation is depicted as a yellow sphere. **b** tRNA is modeled on the basis of the SAXS data for the TcdA-tRNA^{Lys}(UUU) complex bound to the two outer rims of TcdA (represented as in **a**), where most of the positively charged surface is found. The two tRNA molecules bind to spatially separated and independent surface patches in a symmetric arrangement. The modified ct⁶A37 nucleotide is shown in cyan. **c** Detailed view of the TcdA-tRNA binding mode. TcdA is represented as in (**a**), with Cys66 sulfur atom shown as a green sphere. The tRNA molecule on the front inserts its anticodon-stem loop into the ATP-binding pocket (in spheres and CPK colors), with ct⁶A37 (cyan) facing the catalytic site.

doi:10.1371/journal.pone.0118606.g009

Furthermore, the estimation of the molecular weight using either the SAXS data (MW_{SAXS}) or the AUC (MW_{AUC}) data supports that the TcdA-tRNA^{Lys}(UUU) complex contains a TcdA dimer and two tRNA molecules in solution (MW_{SAXS} : 90.0 kDa; MW_{AUC} : 114 kDa; compare with the theoretical MW for a 2:2 complex of 104 kDa) (S2 Table). In addition, SAXS data were collected also for TcdA, validating the crystallographic structure and permitting a more direct comparison of the complex to its components (S1 and S2 Tables).

Ab initio shape restoration from the SAXS scattering data was then used to generate the molecular envelope for the complex (Fig. 8B) and the free tRNA (Fig. 8C) with DAMMIF [45] and MONSA [53,54]. MONSA performs multiphase bead modeling for complexes with different contrasts but requires that the component structures do not undergo conformational changes in the complex. Preliminary analysis of the restored shapes suggested that conformational changes in TcdA were likely to be necessary for productive binding, therefore we used DAMMIF for all subsequent analysis. Since the TcdA-tRNA^{Lys}(UUU) is two-fold symmetric, we generated bead models imposing either *P1* or *P2* symmetry. In both cases the resultant models were similar, but the shapes obtained with *P2* symmetry reproduced better the symmetry known to be present in TcdA homodimer and were kept for further analysis. The overall shape of the TcdA-tRNA^{Lys}(UUU) retrieved by SAXS contains a central body accounting for about 75% of the volume and two symmetric protrusions that extend to either side normal to the symmetry axis that we identified as tRNA^{Lys}(UUU) (Fig. 8B).

To gain insight into the structure of TcdA-tRNA^{Lys}(UUU) we attempted to model the complex into the SAXS envelope, using the structure of TcdA-ATP and a representative structure of tRNA (PDB 4JXX). The shape and volume of the SAXS envelope indicates that TcdA could occupy a central location in the complex with two tRNA molecules symmetrically bound on either side of the complex (Fig. 8B), involving extensive contacts with the positively charged surface of TcdA (Fig. 9). Sequential or simultaneous rigid body fitting of TcdA-ATP and tRNA was possible with minimal intermolecular clashes, achieving a good description of the

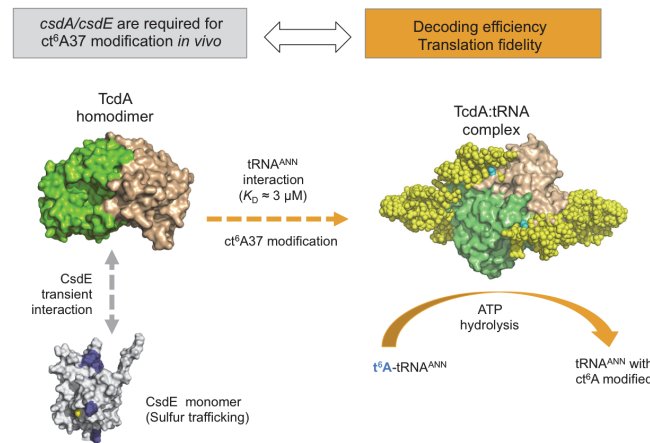


Fig 10. Interaction network for ct^6A37 -tRNA^{ANN} modification. TcdA (surface representation, monomer chains in green and wheat colors) interacts transiently but specifically with the sulfur acceptor CsdE (in grey, with TcdA-binding surface patches in blue and Cys61 in yellow), linking with the CsdA-CsdE cysteine desulfurase system and sulfur trafficking, which are known to be required for ct^6A37 synthesis *in vivo* [16]. The CsdE-TcdA transient interaction is represented by a grey double-headed arrow. TcdA interacts with tRNA^{ANN} (K_D in the μM range) in a 2:2 complex that harbors ATP-dependent t⁶A37 dehydratase activity. The ct^6A37 hypermodification is important for decoding efficiency and translation fidelity.

doi:10.1371/journal.pone.0118606.g010

data (χ^2 2.3; R_{SAS} 0.001) (Fig. 8B). The SAXS-based model of TcdA-tRNA complex that emerges (Fig. 8 and 9) reveals how the flat surface beneath the dimer interface and the surface motifs involved in organizing the K^+ binding site, which have been extensively modified from the E1-like enzyme primitive fold, play crucial roles in binding tRNA. This flattened interface provides an ideal docking platform for the flat side of the tRNA molecules while guiding the stem loop to position A37 in close proximity to the TcdA active site, and its electrostatic properties, complementary to those of tRNA, provide substantial binding stabilization. The large and symmetric electropositive surface patches on the TcdA structure located at the outer rims of the complex, which extend up and including the ATP binding site, provide an extensive positively charge surface for tRNA binding. Projecting the electrostatic potential surface derived from the crystallographic structures onto the shapes restored by SAXS further confirms that the TcdA surfaces involved in tRNA binding are also the most electropositive surface patches in the complex. It is then conceivable that a ternary TcdA-tRNA^{ANN} complex may form as an intermediate toward ct^6A37 modification from which ATP hydrolysis and t⁶A37 cyclization steps may take place subsequently.

Although accurate modeling of the tRNA modification reaction within the TcdA-tRNA complex is impeded by the lack of high-resolution structures for the complex, the SAXS-based model of the complex is compatible with the known biological function of TcdA, whereby the modified t⁶A37 nucleotide in the ASL would fit snugly in the active site groove of TcdA (Fig. 9C), in close proximity to active-site residues and ATP. Although very little is known about how TcdA cyclizes t⁶A37, it is safe to assume that the N⁶-threonylcarbamoyl side chain of t⁶A37 will have to reach inside the ATP-binding pocket for the reaction to proceed, since ATP hydrolysis has been shown to occur concomitantly with reaction turnover. Further research will be necessary to elucidate the mechanistic details of the interaction, including the relative positions of the TcdA catalytic residues, the ATP substrate, and the t⁶A37 side chain.

Interestingly, the TcdA loop that is not defined in our crystal structures (residues 214–236) and that extends out of the K^+ binding pocket right underneath the ATP binding site, would be well placed to interact with the tRNA molecules. It is currently unknown whether this loop will play a predominantly stabilizing, catalytic, or an intermediate, role. Since the loop sequence contains several positively charged residues and two potentially catalytic cysteine residues (Cys220 and Cys234), it could perform complex tasks both securing tRNA binding in a productive orientation and, perhaps, collaborating in the cyclization of t^6A37 . More extensive conformational changes at the TcdA-tRNA interface cannot be ruled out. For example, the largest crystallographic *B*-factors pertain to residues in helix $\alpha 7$ and the flap-like loop between $\alpha 7$ - $\beta 7$, facing toward the hypothetical tRNA interface. Those motifs could rearrange upon binding to a conformation where helix $\alpha 6$ and neighboring motifs could present shape and electrostatic properties resulting in an even more complementary binding surface for tRNA. A crystal structure of the TcdA-tRNA complex would shed light on these questions thereby allowing a more precise modeling of the catalytic mechanism for the TcdA-mediated biosynthesis of ct^6A .

Conclusions

The crystal structures of MoeB/E1-like TcdA in complex with ATP and AMP and the SAXS-based bead models of the TcdA-tRNA^{Lys}(UUU) reveal the basis for the association between TcdA and tRNA^{ANN}. The TcdA-tRNA^{Lys}(UUU) is a 2:2 complex in which two tRNA molecules bind independently to positively charged surfaces encompassing both TcdA chains. In addition, TcdA can interact transiently with the sulfur-acceptor CsdE through a specific surface patch in the vicinity to the catalytic Cys61. These results provide a molecular basis for understanding the tRNA hypermodification function of TcdA that ensures decoding efficiency and translation fidelity of tRNA^{ANN} by the ribosome, and its connection with the sulfur trafficking systems that support TcdA function *in vivo* (Fig. 10).

Supporting Information

S1 Fig. Structure of t^6A37 in the anticodon stem loop (ASL) of tRNA^{Lys}(UUU).
(PDF)

S2 Fig. Guinier plots. (a) *E. coli* TcdA-ATP-Mg²⁺, (b) free tRNA^{Lys}(UUU), (c) *E. coli* TcdA-ATP-Mg²⁺ + tRNA^{Lys}(UUU), and (d) BMOE cross-linked *E. coli* TcdA-CsdE complex.
(PDF)

S3 Fig. Cys66 is located in the active site of TcdA.
(PDF)

S4 Fig. Structural comparison of TcdA and MoeB.
(PDF)

S5 Fig. SAXS shape of BMOE-crosslinked CsdE-TcdA complex.
(PDF)

S1 Table. Summary of SAXS invariant parameters as inferred from data processing and analysis with ScÅtter, Primus, and GNOM.
(PDF)

S2 Table. Comparison of molecular mass estimates calculated from SAXS and SE-AUC.
(PDF)

S1 Text. Analysis of SAXS data for the BMOE cross-linked TcdA-CsdE complex.
(PDF)

Acknowledgments

We gratefully acknowledge access to the ALBA Synchrotron Light Source, Barcelona, Spain, for provision of synchrotron radiation facilities at the BL13-XALOC beamline; Synchrotron SOLEIL, Paris, France, at the PROXIMA 2A beamline; and the European Synchrotron Radiation Facility, Grenoble, France, at the BM29 BioSAXS beamline; and the staff for excellent support at the three synchrotrons. The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under BioStruct-X (grant agreement N° 283570). M.L.E. acknowledges the support of the Ph.D. program in in Biochemistry, Molecular Biology and Biomedicine of the *Universidad Complutense de Madrid* (UCM).

Author Contributions

Conceived and designed the experiments: FJF JJB MCV. Performed the experiments: MLE AA MS AR WS MB FJF MCV. Analyzed the data: MLE AA MS AR WS FJF JJB MCV. Contributed reagents/materials/analysis tools: MLE AA MS AR WS MB FJF JJB MCV. Wrote the paper: MLE AA AR MC FJF JJB MCV.

References

1. Grosjean H. Nucleic Acids Are Not Boring Long Polymers of Only Four Types of Nucleotides: A Guided Tour. In: Grosjean H, editor. *DNA and RNA Modification Enzymes: Structure, Mechanism, Function and Evolution*. Austin: Landes Bioscience; 2000. pp. 1–18.
2. Machnicka MA, Milanowska K, Osman Oglou O, Purta E, Kurkowska M, Olchowik A, et al. MODO-MICS: a database of RNA modification pathways—2013 update. *Nucleic Acids Res*. 2013; 41: D262–267. doi: [10.1093/nar/gks1007](https://doi.org/10.1093/nar/gks1007) PMID: [23118484](https://pubmed.ncbi.nlm.nih.gov/23118484/)
3. Suzuki T. Biosynthesis and function of tRNA wobble modifications. In: Grosjean H, editor. *Fine-Tuning of RNA Functions by Modification and Editing*. Springer Berlin Heidelberg; 2005. pp. 23–69.
4. Agris PF, Vendeix FA, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. *J Mol Biol*. 2007; 366: 1–13. PMID: [17187822](https://pubmed.ncbi.nlm.nih.gov/17187822/)
5. El Yacoubi B, Bailly M, de Crecy-Lagard V. Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu Rev Genet*. 2012; 46: 69–95. doi: [10.1146/annurev-genet-110711-155641](https://doi.org/10.1146/annurev-genet-110711-155641) PMID: [22905870](https://pubmed.ncbi.nlm.nih.gov/22905870/)
6. Schweizer MP, Chheda GB, Baczynskyj L, Hall RH. Aminoacyl nucleosides. VII. N-(Purin-6-ylcarbamoyl)threonine. A new component of transfer ribonucleic acid. *Biochemistry*. 1969; 8(8): 3283–3289. PMID: [4897334](https://pubmed.ncbi.nlm.nih.gov/4897334/)
7. Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, et al. Deficit of tRNA(Lys) modification by Cdkal1 causes the development of type 2 diabetes in mice. *J Clin Invest*. 2011; 121: 3598–3608. doi: [10.1172/JCI58056](https://doi.org/10.1172/JCI58056) PMID: [21841312](https://pubmed.ncbi.nlm.nih.gov/21841312/)
8. Niimi T, Niimi O, Nureki T, Yokogawa N, Hayashi K, Nishikawa K, et al. Recognition of the Anticodon Loop of tRNA Ile 1 by Isoleucyl-tRNA Synthetase from *Escherichia coli*. *Nucleosides & Nucleotides*. 1994; 13: 1231–1237.
9. Yarian C, Townsend H, Czystkowski W, Sochacka E, Malkiewicz AJ, Guenther R, et al. Accurate translation of the genetic code depends on tRNA modified nucleosides. *J Biol Chem*. 2002; 277(19): 16391–16395. PMID: [11861649](https://pubmed.ncbi.nlm.nih.gov/11861649/)
10. Weissenbach J, Grosjean H. Effect of threonylcarbamoyl modification (t6A) in yeast tRNA Arg III on codon-anticodon and anticodon-anticodon interactions. A thermodynamic and kinetic evaluation. *Eur J Biochem*. 1981; 116: 207–213. PMID: [6788546](https://pubmed.ncbi.nlm.nih.gov/6788546/)
11. Phelps SS, M A., Agris PF, Joseph S. Modified nucleotides in tRNA(Lys) and tRNA(Val) are important for translocation. *J Mol Biol*. 2004; 338: 439–444. PMID: [15081802](https://pubmed.ncbi.nlm.nih.gov/15081802/)
12. Lin CA, Ellis SR, True HL. The Sua5 protein is essential for normal translational regulation in yeast. *Mol Cell Biol*. 2010; 30: 354–363. doi: [10.1128/MCB.00754-09](https://doi.org/10.1128/MCB.00754-09) PMID: [19884342](https://pubmed.ncbi.nlm.nih.gov/19884342/)
13. Stuart JW, Gdaniec Z, Guenther R, Marszalek M, Sochacka E, Malkiewicz A, et al. Functional anticodon architecture of human tRNA^{Lys3} includes disruption of intraloop hydrogen bonding by the naturally occurring amino acid modification, t6A. *Biochemistry*. 2000; 39: 13396–13404. PMID: [11063577](https://pubmed.ncbi.nlm.nih.gov/11063577/)

14. Sundaram M, Durant PC, Davis DR. Hypermodified nucleosides in the anticodon of tRNA(Lys) stabilize a canonical U-turn structure. *Biochemistry*. 2000; 39: 15652. PMID: [11112555](#)
15. Murphy FVt, Ramakrishnan V, Malkiewicz A, Agris PF. The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat Struct Mol Biol*. 2004; 11: 1186–1191. PMID: [15558052](#)
16. Miyauchi K, Kimura S, Suzuki T. A cyclic form of N6-threonylcarbamoyladenosine as a widely distributed tRNA hypermodification. *Nat Chem Biol*. 2013; 9: 105–111. doi: [10.1038/nchembio.1137](#) PMID: [23242255](#)
17. Ikeuchi Y, Shigi N, Kato J-I, Nishimura A, Suzuki T. Mechanistic insights into sulfur relay by multiple sulfur mediators involved in thioridine biosynthesis at tRNA wobble positions. *Mol Cell*. 2006; 21: 97–108. PMID: [16387657](#)
18. Trotter V, Vinella D, Loiseau L, Ollagnier de Choudens S, Fontecave M, Barras F. The CsdA cysteine desulphurase promotes Fe/S biogenesis by recruiting Suf components and participates to a new sulphur transfer pathway by recruiting CsdL (ex-YgdL), a ubiquitin-modifying-like protein. *Mol Microbiol*. 2009; 74: 1527–1542. PMID: [20054882](#)
19. Bolstad HM, Botelho DJ, Wood MJ. Proteomic analysis of protein-protein interactions within the Cysteine Sulfinate Desulfurase Fe-S cluster biogenesis system. *J Proteome Res*. 2010; 9: 5358–5369. doi: [10.1021/pr1006087](#) PMID: [20734996](#)
20. Loiseau L, Ollagnier-de Choudens S, Lascoux D, Forest E, Fontecave M, Barras F. Analysis of the heteromeric CsdA-CsdE cysteine desulfurase, assisting Fe-S cluster biogenesis in *Escherichia coli*. *J Biol Chem*. 2005; 280: 26760–26769. PMID: [15901727](#)
21. Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J*. 2000; 78: 1606–1619. PMID: [10692345](#)
22. Cole JL. Analysis of heterogeneous interactions. *Methods Enzymol*. 2004; 384: 212–232. PMID: [15081689](#)
23. Ucci JW, Cole JL. Global analysis of non-specific protein-nucleic interactions by sedimentation equilibrium. *Biophys Chem*. 2004; 108: 127–140. PMID: [15043926](#)
24. Laue TM, Shah BD, Ridgeway TM, Pelletier SL. Computer-aided interpretation of analytical sedimentation data for proteins. *Anal Ultracentrifugation Biochem Polym Sci*; 1992. pp. 90–125.
25. Kim S, Kim KY, Park JK, Lee BI, Kim YG, Park S. Overproduction, crystallization and preliminary X-ray crystallographic analysis of *Escherichia coli* tRNA N(6)-threonylcarbamoyladenosine dehydratase. *Acta Crystallogr F Struct Biol Commun*. 2014; 70: 1517–1520. doi: [10.1107/S2053230X14020883](#) PMID: [25372820](#)
26. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr*. 2010; 66: 125–132. doi: [10.1107/S0907444909047337](#) PMID: [20124692](#)
27. Evans PR, Murshudov GN. How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr*. 1998; 54: 1274–1282. doi: [10.1107/S0907444913000061](#) PMID: [23793146](#)
28. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr*. 2011; 67: 235–242. doi: [10.1107/S0907444910045749](#) PMID: [21460441](#)
29. Lake MW, Wuebbens MM, Rajagopalan KV, Schindelin H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature*. 2001; 414: 325–329. PMID: [11713534](#)
30. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. *J Appl Crystallogr*. 2007; 40: 658–674. PMID: [19461840](#)
31. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010; 66: 486–501. doi: [10.1107/S0907444910007493](#) PMID: [20383002](#)
32. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*. 2012; 68: 352–367. doi: [10.1107/S0907444912001308](#) PMID: [22505256](#)
33. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*. 2010; 66: 12–21. doi: [10.1107/S0907444909042073](#) PMID: [20057044](#)
34. van de Weert M. Fluorescence quenching to study protein-ligand binding: common errors. *J Fluoresc*. 2010; 20: 625–629. doi: [10.1007/s10895-009-0572-x](#) PMID: [19997966](#)
35. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins*. 2005; 59: 687–696. PMID: [15815974](#)
36. Grzesiek S, Bax A. An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins. *J Magn Reson*. 1992; 99: 201–207.

37. Grzesiek S, Bax A. Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR. *J Am Chem Soc.* 1992; 114: 6291–6293.
38. Liu G, Li Z, Chiang Y, Acton T, Montelione GT, Murray D, et al. High-quality homology models derived from NMR and X-ray structures of *E. coli* proteins YgdK and Suf E suggest that all members of the YgdK/Suf E protein family are enhancers of cysteine desulfurases. *Protein Sci.* 2005; 14: 1597–1608. PMID: [15930006](#)
39. Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, et al. Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. *J Synchrotron Rad.* 2013; 20: 660–664.
40. Round A, Brown E, Marcellin R, Kapp U, Westfall CS, Jez JM, et al. Determination of the GH3.12 protein conformation through HPLC-integrated SAXS measurements combined with X-ray crystallography. *Acta Crystallogr D Biol Crystallogr.* 2013; 69: 2072–2080. doi: [10.1107/S0907444913019276](#) PMID: [24100325](#)
41. Guinier A, Fournet G (1955) *Small-Angle Scattering of X-Rays*. New York, NY: John Wiley & Sons.
42. Porod G. General theory. In: Glatter O, Kratky O, editors. *Small Angle X-Ray Scattering*. New York, NY: Academic Press; 1982. pp. 17–51.
43. Svergun DI. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Cryst.* 1992; 25: 495–503.
44. Svergun DI. CRY SOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Cryst.* 1995; 28: 768–773.
45. Franke D, Svergun DI. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Cryst.* 2009; 42: 342–346.
46. Volkov VV, Svergun DI. Uniqueness of ab initio shape determination in small-angle scattering. *J Appl Cryst.* 2003; 36: 860–864.
47. Rambo RP, Tainer JA. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature.* 2013; 496: 477–481. doi: [10.1038/nature12070](#) PMID: [23619693](#)
48. Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, et al. The Pfam protein families database. *Nucleic Acids Res.* 2012; 40: D290–301. doi: [10.1093/nar/gkr1065](#) PMID: [22127870](#)
49. Krissinel E. Macromolecular complexes in crystals and solutions. *Acta Crystallogr D Biol Crystallogr.* 2011; 67: 376–385. doi: [10.1107/S0907444911007232](#) PMID: [21460456](#)
50. Walker JE, Saraste M, Runswick MJ, Gay NJ. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* 1982; 1: 945–951. PMID: [6329717](#)
51. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr.* 2004; 60: 2256–2268. PMID: [15572779](#)
52. Kim S, Park S. Structural Changes during Cysteine Desulfurase CsdA and Sulfur Acceptor CsdE Interactions Provide Insight into the trans-Persulfuration. *J Biol Chem.* 2013; 288: 27172–27180. doi: [10.1074/jbc.M113.480277](#) PMID: [23913692](#)
53. Svergun DI. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J.* 1999; 76: 2879–2886. PMID: [10354416](#)
54. Svergun DI, Nierhaus KH. A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. *J Biol Chem.* 2000; 275: 14432–14439. PMID: [10799526](#)
55. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A.* 2001; 98: 10037–10041. PMID: [11517324](#)
56. Schrodinger, LLC (2010) The PyMOL Molecular Graphics System, Version 1.3r1.

Mechanism of Sulfur Transfer Across Protein–Protein Interfaces: The Cysteine Desulfurase Model System

Francisco J. Fernández,[†] Ana Ardá,^{†,‡,●} Miguel López-Esteva,[†] Juan Aranda,[§] Esther Peña-Soler,[†] Fernando Garces,^{||} Adam Round,^{⊥,‡} Ramón Campos-Olivas,[○] Marta Bruix,[▲] Miquel Coll,^{¶,□} Iñaki Tuñón,[§] Jesús Jiménez-Barbero,^{†,‡,×,∞,●} and M. Cristina Vega^{*,†}

[†]Chemical and Physical Biology Department, Center for Biological Research (CIB-CSIC), 28040 Madrid, Spain

[‡]CIC bioGUNE, Bizkaia Technology Park, Building 801A, 48170 Derio, Spain

[×]Ikerbasque, Basque Foundation for Science, Maria Diaz de Haro 13, 48009 Bilbao, Spain

[∞]Department of Organic Chemistry II, Faculty of Science & Technology, University of the Basque Country, 48940 Leioa, Bizkaia, Spain

[§]Departamento de Química Física, Universitat de València, 46100 Burjassot, Spain

^{||}The Scripps Research Institute, La Jolla, 92037 California, United States

[⊥]European Molecular Biology Laboratory, Grenoble Outstation, 38042 Grenoble, France

[#]Unit for Virus Host-Cell Interactions, University Grenoble Alpes-EMBL-CNRS, 38042 Grenoble, France

[○]Spanish National Cancer Research Center (CNIO), 28029 Madrid, Spain

[▲]Instituto de Química Física Rocasolano (IQFR-CSIC), 28006 Madrid, Spain

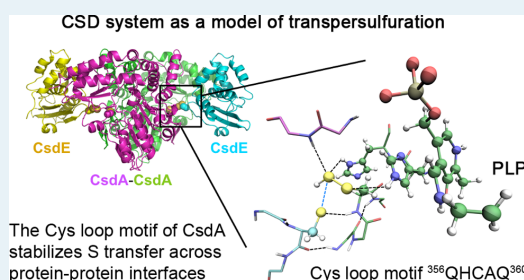
[¶]Institut de Biologia Molecular de Barcelona (IBMB-CSIC), 08028 Barcelona, Spain

[□]Institute for Biomedical Research (IRB Barcelona), 08028 Barcelona, Spain

Supporting Information

ABSTRACT: CsdA cysteine desulfurase (the sulfur donor) and the CsdE sulfur acceptor are involved in biological sulfur trafficking and in iron–sulfur cluster assembly in the model bacterium *Escherichia coli*. CsdA and CsdE form a stable complex through a polar interface that includes CsdA Cys328 and CsdE Cys61, the two residues known to be involved in the sulfur transfer reaction. Although mechanisms for the transfer of a sulfur moiety across protein–protein interfaces have been proposed based on the IscS–IscU and IscS–TusA structures, the flexibility of the catalytic cysteine loops involved has precluded a high resolution view of the active-site geometry and chemical environment for sulfur transfer. Here, we have used a combination of X-ray crystallography, solution NMR and SAXS, isothermal calorimetry, and computational chemistry methods to unravel how CsdA provides a specific recognition platform for CsdE and how their complex organizes a composite functional reaction environment. The X-ray structures of persulfurated (CsdA)₂ and persulfurated (CsdA–CsdE)₂ complexes reveal the crucial roles of the conserved active-site cysteine loop and additional catalytic residues in supporting the transsulfuration reaction. A mechanistic view of sulfur transfer across protein–protein interfaces that underpins the requirement for the conserved cysteine loop to provide electrostatic stabilization for the in-transfer sulfur atom emerges from the analysis of the persulfurated (CsdA–CsdE)₂ complex structure.

KEYWORDS: CsdA–CsdE system, cysteine desulfurase, sulfur trafficking, computational chemistry, transsulfuration, biogenesis



INTRODUCTION

Understanding key molecular recognition processes at atomic level resolution is one of the essential topics at the frontier between chemistry and biology. Cellular mobilization of elemental sulfur (S) from L-Cys is a crucial step for the biosynthesis of several vital entities, including Fe–S clusters, thiolated amino acids and vitamins (thiamine, biotin), lipoic

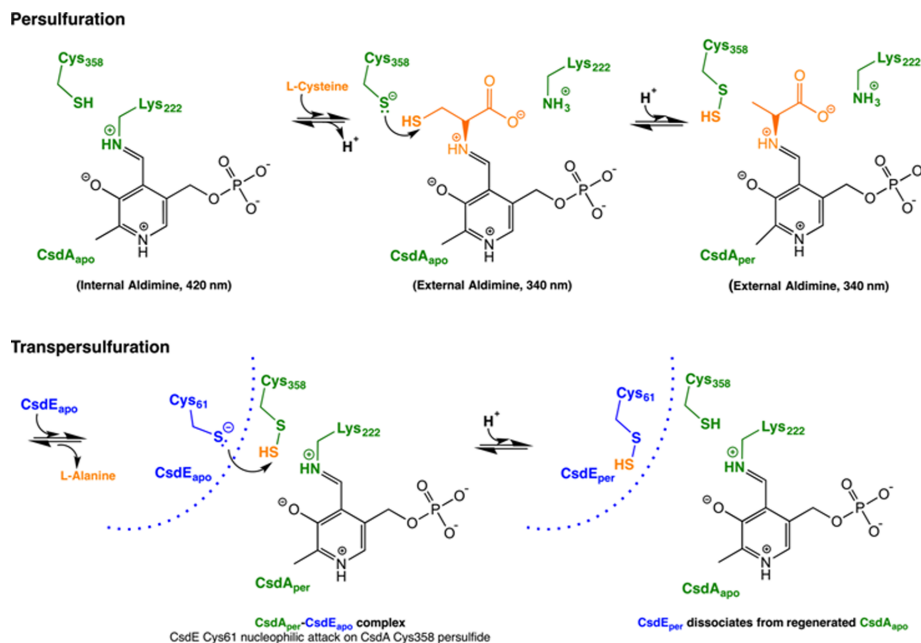
acid, protein cofactors (molybdenum cofactor), and various modified RNA nucleosides.^{1,2} Trafficking and deployment of S for these biosynthetic pathways is accomplished by an intricate

Received: February 3, 2016

Revised: April 29, 2016

Published: May 5, 2016

Scheme 1. Generic Enzymatic Mechanism Proposed to Explain the Persulfuration of Cysteine Desulfurases and the Transpersulfuration of Cosubstrate Sulfur Acceptors



network of conserved multiprotein complexes, which act as donors and acceptors of S via a reactive persulfide.^{3,4} This exchange of S across protein–protein interfaces, or transpersulfuration, fulfills a range of important biological functions: it circumvents the low solubility of sulfane S, avoids the liberation of the toxic persulfide anion into the cytosol, and, importantly, allows for a more specific S transfer between proteins.

In *Escherichia coli* (*Ec*) three distinct systems can initiate sulfur transfer pathways: ISC (iron–sulfur clusters), SUF (sulfur mobilization), and CSD (cysteine desulfurase).^{2,5–7} The ISC system comprises housekeeping proteins that are required for the constitutive assembly of Fe–S clusters and for the biosynthesis of several sulfur-containing cofactors and tRNAs. The SUF system is upregulated during O₂ stress and Fe starvation.⁸ The CSD system has been implicated so far in the reconstitution of the [4Fe–4S] NadA holoenzyme,⁹ the supply of S atoms for Fe–S cluster assembly via the SUF pathway, and in the biosynthesis of an unidentified sulfur-containing molecule.¹⁰ More recently, the CSD system has emerged as the likely sulfur relay system dedicated to the synthesis of cyclic N'-threonylcarbamoyladenosine (ct⁶A), a nonthiolated tRNA hypermodification found in bacteria, protists, fungi, and plants.^{11,12} In contrast to the ISC system, the mechanism of transpersulfuration in the SUF and CSD systems⁹ is less well understood.

Sulfur transfer is initiated in all known systems by the persulfuration of an active site Cys residue by a pyridoxal 5'-phosphate (PLP) dependent cysteine desulfurase (E.C. 2.8.1.7),¹³ IscS (ISC),¹⁴ SufS (previously known as CsdB) (SUF),¹⁵ and CsdA (CSD).⁹ The resultant persulfide nucleophile intermediate¹⁶ is responsible for delivering the terminal S_γ atom of the persulfide to the thiol group of another

reactive Cys residue in a cosubstrate S acceptor protein, through a so-called transpersulfuration reaction (Scheme 1). It remains an open question whether the mechanisms of transpersulfuration across different cysteine desulfurases/S acceptors are as well conserved as those of the intramolecular persulfuration reaction, given the variety of interaction modes that have thus far been described. Despite CsdA is able to catalyze the S transfer from L-Cys to Cys358, the presence of CsdE enhances the rate of the reaction 2-fold.⁹ Also some degree of promiscuity is observed in CsdA, and cross-talk events between the CSD and SUF systems have been detected, increasing the cellular robustness under changing environmental conditions^{9,10} (e.g., oxidative stress).

Despite their shared enzymatic mechanism, cysteine desulfurases are classified in two groups based on sequence and structural homology: group I (e.g., IscS) and group II (e.g., CsdA/SufS).¹³ In particular, group I and II cysteine desulfurases differ in the sequence of a conserved peptide motif enclosing the catalytic Cys residue (the Cys loop,¹³ ³⁵⁶QHCAQ³⁶⁰ in CsdA), whose functional relevance has hitherto remained enigmatic. The Cys loop of group I cysteine desulfurases is characterized by being highly mobile, as highlighted in the recent crystallographic studies of IscS in complex with IscU and TusA, where the mobility of the Cys loop makes the transpersulfuration possible despite the markedly different surfaces employed by IscS in these two complexes.^{17,18} The crystal structure of (Fe₂S₂–(IscS–IscUD3SA))₂ complex from *Archaeoglobus fulgidus* VC-16 (Af) shows an ordered Cys loop only when the Fe–S cluster is being assembled.¹⁸ In contrast to IscS, the Cys loop of desulfurases from SUF and CSD systems (SufS/CsdA group II cysteine desulfurases),^{19,20} maintains an essentially rigid conformation. The crystal structure of (CsdA_{apo}–CsdE_{apo})₂

complex²⁰ revealed that the Cys-loop remained unchanged upon CsdE binding. In this case, CsdE undergoes conformational changes in H7 that leads to a better fit into the active site groove of CsdA.²⁰ Critically, the catalytic residue Cys61 from CsdE in the (CsdA_{apo}–CsdE_{apo})₂ complex is disordered and entirely missing in the electron density map, thus limiting the chemical inferences that can be made of the interprotein S transfer. As a result, important questions such as how the chemical environment in the active site stabilizes the persulfide on its journey through different protein–protein complexes (in the CSD system or during CSD–SUF crosstalk), why the persulfuration rate is enhanced in the presence of CsdE, or how sensitive the chemistry of transpersulfuration is to oxidative stress remain still unanswered.

In order to shed light into the process of transpersulfuration across the interface between (CsdA)₂ and CsdE, we applied biophysical, structural biology, and chemical computational approaches. We have solved the crystal structure of (CsdA)₂ in different persulfuration and oxidative states, and the active (persulfurated) complex of CsdA and CsdE. Furthermore, we have analyzed the CsdA–CsdE interaction in solution by NMR and compared the stability of the complexes of CsdA and CsdE and SufE with isothermal titration calorimetry (ITC). And finally, with the aid of QM/MM calculations, we propose a mechanism for the transpersulfuration reaction that is shared across group II Cys desulfurases.

EXPERIMENTAL PROCEDURES

Expression and Purification of Proteins. The detailed procedures used to prepare (CsdA)₂ in various states including persulfurated (CsdA_{per})₂ and S-sulfated (CsdA_s)₂, CsdE, and SufE are presented in the [Supporting Information](#) (SI text).

Reconstitution and Purification of Unmodified and Persulfurated (CsdA–CsdE)₂. The unmodified (CsdA–CsdE)₂ complex was prepared by mixing purified components, (CsdA)₂ (obtained from His₆-CsdA after proteolytic removal of the histidine tag with TEV protease) and CsdE–His₆, at a 1:1.5 molar ratio for 30 min at 277 K, and subjecting the mixture to gel filtration chromatography on a Superdex 200 10/300GL (GE Healthcare) in 10 mM HEPES pH 7.4, 50 mM NaCl, 0.1 mM EDTA. Peak fractions were concentrated to 20 mg/mL using 10-kDa ultrafiltration devices (Millipore) and stored frozen at 193 K. The persulfurated (CsdA_{per}–CsdE_{per})₂ complex was prepared by incubating the unmodified (CsdA–CsdE)₂ complex with 5 mM L-cysteine for 30 min at 298 K and then separating the persulfurated (CsdA_{per}–CsdE_{per})₂ complex from excess CsdE–His₆, L-cysteine, and L-alanine, by gel filtration (as indicated above). The (CsdA_{per}–CsdE_{per})₂ complex was concentrated and stored as described for the unmodified (CsdA–CsdE)₂ complex.

Crystallization and Structure Determination of (CsdA)₂ Variants. Crystallization, X-ray data collection, structure determination, and refinement of (CsdA)₂, persulfurated (CsdA_{per})₂ (Figure S1), and S-sulfated (CsdA_s)₂, are described in detail in the [SI](#). Complete data processing and refinement statistics are given in [Tables 1](#) and [S1](#).

Crystallization of Persulfurated (CsdA–CsdE)₂. Persulfurated (CsdA_{per}–CsdE_{per})₂ complex was crystallized by sitting-drop vapor diffusion using 1 μ L complex and 1 μ L reservoir condition (0.1 M Tris–HCl (pH 8.5), 0.1 M magnesium acetate, 10% (w/v) PEG 8000) at 293 K. Prismatic crystals with an intense yellow color developed in 1–4 d, reaching final dimensions of $\sim 80 \mu\text{m} \times 20 \mu\text{m} \times 20 \mu\text{m}$. Prior to diffraction

Table 1. X-ray Data Collection and Refinement Statistics

	(CsdA _{per}) ₂	(CsdA _{per} –CsdE _{per}) ₂
wavelength (Å)	0.87260	0.98011
resolution (Å)	36.7–2.38 (2.47–2.38)	48.3–2.49 (2.63–2.49)
space group	<i>P</i> ₂ ₁ ₂ ₁	<i>P</i> ₂ ₁ ₂ ₁
unit cell (Å, deg)	58.3 99.5 141.7 90 90 90	74.8 115.1 604.9 90 90 90
total reflections	124 259 (12 453)	625 506 (64 000)
unique reflections	29 943 (3051)	178 722 (23 193)
multiplicity	4.1 (4.1)	3.5 (2.8)
completeness (%)	89.40 (91.51)	97.90 (88.30)
mean <i>I</i> / σ (<i>I</i>)	11.13 (2.40)	7.40 (1.80)
Wilson <i>B</i> -factor (Å ²)	22.65	49.53
<i>R</i> _{merge}	0.176 (0.736)	0.127 (0.553)
<i>R</i> _{meas}	0.1966	0.1490
<i>CC</i> _{1/2}	0.990 (0.703)	0.989 (0.581)
<i>CC</i> *	0.997 (0.909)	0.996 (0.815)
<i>R</i> _{work}	0.1669 (0.2306)	0.2532 (0.3566)
<i>R</i> _{free}	0.2245 (0.3001)	0.2768 (0.4141)
no. of non-H atoms	6525	34 325
macromolecules	6139	32 886
ligands	87	418
water	299	1021
protein residues	800	4296
RMSD (bonds) (Å)	0.004	0.010
RMSD (angles) (deg)	0.77	1.20
Ramachandran favored (%)	97.00	96.00
Ramachandran outliers (%)	0.00	0.68
Clashscore	8.29	8.09
average <i>B</i> -factor	30.50	68.70
macromolecules	29.80	69.70
ligands	62.10	59.30
solvent	33.70	40.40
PDB ID	SFT5	SFT8

data collection, crystals were equilibrated in crystallization condition and 30% (v/v) glycerol, then mounted in nylon loops and flash-frozen in liquid nitrogen.

X-ray Data Collection. A complete data set consisting of 720 images at 0.5°-oscillation angle (360° in total) was collected on ID23-1 beamline (ESRF, Grenoble, France), equipped with an ADSC Q310r detector, at 100 K and using a crystal-to-detector distance of 253 mm. Persulfurated (CsdA_{per}–CsdE_{per})₂ crystals belonged to the *P*₂₁₂₁ orthorhombic space group with cell dimensions *a* = 74.86 Å, *b* = 115.14 Å, and *c* = 604.97 Å, and diffracted to 2.49-Å resolution. The asymmetric unit contained four biological heterotetramers, thereby yielding eight independent CsdA_{per}–CsdE_{per} heterodimers/interfaces. The data set was processed with XDS²¹ and scaled with Aimless.²² Complete data processing statistics are given in [Table 1](#).

Structure Determination and Refinement. The crystal structures of (CsdA_{per})₂ at 1.76 Å resolution (*R*/*R*_{free} values of 0.166/0.224) ([Table 1](#) and [Figure S1](#)) and S-sulfated (CsdA_s)₂ at 2.05 Å resolution (*R*/*R*_{free} values of 0.154/0.190) ([Table S1](#)) were determined by molecular replacement using as search model a structure of a new (CsdA)₂ crystal structure (1.99 Å, *R*/*R*_{free} values of 0.137/0.191) solved independently by

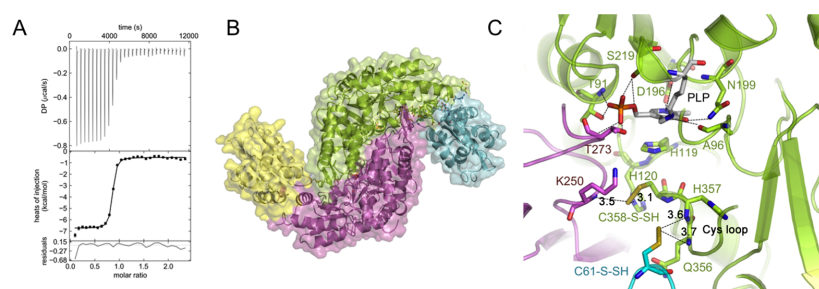


Figure 1. X-ray crystal structure of persulfurated (CsdA–CsdE)₂ complex. (A) ITC thermogram and binding isotherm for the interaction of CsdA with CsdE at 310 K. In the upper panel, peaks indicate the heat released after each consecutive addition of CsdE into the CsdA protein solution. In the lower panel, the data were fit (continuous line) to a single binding model to calculate the thermodynamic parameters. Monomer molar concentrations were used for the fit. (B) Ribbon representation of the persulfurated (CsdA–CsdE)₂ heterotetramer. CsdA chains are shown in green and violet, and CsdE chains in yellow and slate blue. Side chains of PLP and persulfurated CsdA Cys358 and CsdE Cys61 are represented as space-filling models in CPK colors. (C) Structural basis for the conservation of the peptide signature motif of group II cysteine desulfurases (labeled “Cys loop”). Interactions stabilizing the persulfidic S atoms in the CsdA_{per}–CsdE_{per} complex are depicted as dashed lines with key distances shown (Å).

selenomethionine (Se)-SAD (SI text and Table S1). The structure of the persulfurated (CsdA_{per}–CsdE_{per})₂ complex was determined by molecular replacement with PHASER²³ using as search models the crystal structures of (CsdA)₂ (PDB 5FT4) and CsdE (PDB 5FT7) simultaneously during the rotation and translation searches. The molecular replacement solution was then completed manually through cycles of maximum likelihood refinement in phenix.refine²⁴ interspersed with manual building and validation in Coot²⁵ and MolProbity.²⁶ Refinement statistics are in Table 1. Atomic coordinates and structure factors of persulfurated (CsdA_{per}–CsdE_{per})₂ have been deposited with PDB 5FT8.

Protein Expression and Purification of Uniformly Labeled [U-¹³C,¹⁵N]-CsdE-His₆. Uniformly labeled [U-¹³C,¹⁵N]-CsdE-His₆ for nuclear magnetic resonance studies was produced by growing *E. coli* BL21(DE3) bacteria transformed with pET3a-CsdE-His₆ in minimal medium using [U-¹⁵N]-ammonium sulfate and [U-¹³C]-D-glucose as sole sources of nitrogen and carbon, respectively. Purification was carried out as indicated for nonlabeled CsdE-His₆ without modifications.

CsdA-Binding Surface of CsdE-His₆ Assessed by NMR. Proteins used in titration experiments, [U-¹³C,¹⁵N]-CsdE-His₆ and unlabeled CsdA, were dissolved in 50 mM potassium phosphate (pH 7.5), 0.1 mM EDTA, in 90%/10% H₂O/D₂O. NMR experiments were acquired at 298 K on a Bruker Avance 600 MHz spectrometer equipped with a cryoprobe for 1D and 2D, and on a Bruker Avance 800 MHz for 3D. The spectra were processed with TopSpin 2.1 (Bruker), and 2D and 3D spectra were analyzed with the program CCPNmr Analysis.²⁷ 1D ¹H–¹³C and ¹H–¹⁵N HSQC spectra (5120 and 10240 scans respectively), 2D ¹H–¹⁵N HSQC spectra (3072 data points in t₂, 128 in t₁, 96 scans), and 3D spectra (2048 data points in t₃, 30 in t₂ (¹⁵N), and 128 in t₁ (¹³C), 32 scans) were acquired using standard Bruker pulse sequences. 1D heteronuclear spectra were processed with a line broadening of 10 and 20 Hz (¹H–¹³C and ¹H–¹⁵N HSQC respectively). ¹H–¹⁵N backbone resonance assignment for [U-¹³C,¹⁵N]-CsdE-His₆ was performed using HSQC, CBCANH,²⁸ and CBCA(CO)-NH²⁹ spectra. Most NH/Ha crosspeaks could be assigned and were in agreement with published assignments (BMRB code: 5630).³⁰

The interaction between CsdA and [U-¹³C,¹⁵N]-CsdE-His₆ was probed by stepwise addition of unlabeled CsdA to a 0.4 mM [U-¹³C,¹⁵N]-CsdE-His₆. At each titration point (1:0, 1:0.2, 1:0.4, and 1:0.6), ¹H–¹⁵N HSQC spectra were recorded, and the rate of intensity loss was obtained by fitting the volume of every NH/Ha crosspeak at different ratios to an exponential decay function $V = Ae^{-Bx}$, where V is the volume of an NH/Ha crosspeak, x is the CsdE/CsdA molar ratio, and A and B are fitted parameters.

Computational Studies. Protein modeling and quantum mechanical/molecular mechanical (QM/MM) calculations were performed with the programs AMBER,³¹ NAMD,³² fDynamo,³³ and Gaussian09.³⁴ We constructed a model that was derived from the crystal structure of persulfurated (CsdA_{per}–CsdE_{per})₂ complex, with CsdE Cys61 modeled with a thiolate side chain. The QM subsystem consisted of the side chain of the persulfurated CsdA Cys358, the side chains of the CsdE Cys61 and CsdA His119 and His120, and the Lys222-PLP cofactor (69 atoms in total). The QM subsystem was described at the M06-2X/6-31G**³⁵ level of theory while the MM system was described using the AMBER force field as implemented in fDynamo.^{36,37} To saturate the valence of the QM/MM boundary we used the link atom procedure.^{38,39} Additional information on the computational chemistry methodology is presented in the SI.

RESULTS AND DISCUSSION

The interaction between the CsdA homodimer and each CsdE monomer is noncooperative and is more stable than in the complexes (SufS–SufE)₂ or (CsdA–SufE)₂. We performed ITC titration studies to determine the thermodynamic constants of the interaction between CsdA and CsdE (Figure 1). The results showed a thermodynamically (both enthalpically and entropically) favorable curve with an expected 1:1 stoichiometry ($n = 1.072$) and a $K_D \sim 117$ nM ($\Delta G = -9.5$ kcal/mol at 298 K) (Figure 1A, SI text, Table S2, and Figure S2). In contrast to the binding isotherms of IscS–IscU⁴⁰ and SufS–SufE,^{41,42} which are biphasic with exothermic and endothermic phases, the ITC binding isotherm for the CsdA–CsdE interaction is in agreement with a monophasic and exothermic interaction. This difference suggests that the CsdA–CsdE interaction is noncooperative and, in particular,

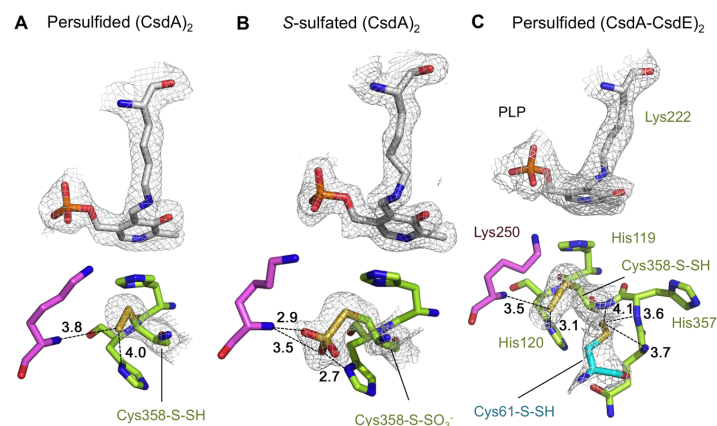


Figure 2. Active site structures of (A) persulfated $(\text{CsdA}_{\text{per}})_2$, (B) S-sulfated $(\text{CsdA}_{\text{ss}})_2$, and (C) persulfated $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$. Amino acid residues are represented in sticks and shown in CPK colors except carbon atoms; carbon atoms are colored in gray (Lys222-PLP), green (CsdA monomer in complex with CsdE), violet (opposite CsdA monomer), and cyan (CsdE). Interacting atoms relevant for the transpersulfuration reaction are connected with black dashed lines labeled with the interaction distances (Å). Electron density maps contoured at 1σ are overlaid on Lys222-PLP and the catalytic Cys residues (CsdA Cys358 and CsdE Cys61).

does not exhibit the flip-flop allosteric regulation mechanism described for other desulfurase-acceptor pairs.

The thermodynamic data obtained indicate that the $(\text{CsdA}-\text{CsdE})_2$ complex exists as a more stable complex than the homologous $(\text{SufS}-\text{SufE})_2$ complex⁴¹ and related complexes, like $(\text{IscS}-\text{IscU})_2$,^{40,43} which have a 10–20-fold weaker affinity. Interestingly, ITC measurements of SufS_{apo} interaction with alkylated SufE_{alk} (prepared by incubation of SufE_{apo} with iodoacetamide, which quantitatively alkylates the reactive Cys51) show a 10-fold tighter K_D (260 nM) than the $\text{SufS}_{\text{apo}}-\text{SufE}_{\text{apo}}$ interaction and a monophasic behavior.⁴¹ The $\text{SufS}_{\text{apo}}-\text{SufE}_{\text{alk}}$ interaction therefore might approximate the $\text{CsdA}-\text{CsdE}$ interaction more closely. Since the peptide containing the alkylated Cys51 in SufE_{alk} is more solvent exposed than the unmodified Cys51, where the same peptide is protected from solvent in a hydrophobic pocket, the structure around Cys51 in SufE_{alk} might adopt a similar conformation to that of CsdE. If this were the case, the tighter K_D for $\text{SufS}_{\text{apo}}-\text{SufE}_{\text{alk}}$ could be the result of the formation of a complex structurally analogous to $\text{CsdA}-\text{CsdE}$.

CsdA is also known to interact with SufE in a crosstalk between the CSD and SUF systems, though less tightly than with CsdE.⁹ We did not detect measurable binding of CsdA and SufE_{apo} by ITC (data not shown), indicating that their association constant is very weak (below mM range). Despite the absence of a stable $(\text{CsdA}-\text{SufE})_2$ complex, CsdA can persulfurate SufE in vitro in a quantitative fashion, indicating that a short-lived catalytic $(\text{CsdA}-\text{SufE})_2$ complex may exist. This observation is compatible with the reported role of CsdA in Fe-S biogenesis by recruiting the SufBCD complex and SufE .¹⁰

Composite Active Site of CsdA Preformed for Persulfuration: Cys358 Sy in $(\text{CsdA}_{\text{per}})_2$ Stabilized by Residues from Both Subunits. To investigate the structural and chemical consequences of the persulfuration of CsdA we solved the crystal structures of persulfated CsdA and of the oxidized S-sulfated CsdA, $(\text{CsdA}_{\text{ss}})_2$, at 1.76 and 2.05 Å resolution, respectively. The two covalently modified CsdA structures showed no significant differences in their fold

arrangement (Figure S1). CsdA belongs to the fold-type I class V aminotransferase or α -family of PLP-dependent enzymes,^{17,40} which are structurally characterized by a constitutive dimer built from identical two-domain subunits (green and violet, Figure 1 A). The covalently modified Cys358 is at the tip of the so-called Cys loop motif in the small domain (residues 4–22 and 294–401) (Figure 1 B and C).¹³ The large domain (residues 23–293) harbors the PLP cofactor and the cysteine-substrate binding pocket. In both crystal structures, easily recognizable electron density indicated the presence of the PLP cofactor as an internal aldimine covalently bound to Lys222 (Figures 1 C and 2).

In $(\text{CsdA}_{\text{per}})_2$, the persulfido moiety on Cys358 is stabilized by hydrogen bonding interactions with key residues, including the imidazole side chain of His119 and the main-chain amide of Lys250* from the other subunit, which also participates in stabilizing the active site during transpersulfuration (Figures 1 C and 2 A). The hydrogen bond network centered on the persulfidic Sδ atom holds the Cys358 modified side chain already in an orientation that is ready for reaction, therefore priming CsdA for transpersulfuration. The dimer interface in CsdA_{per} involves extensive contacts between residues from the large domain and the small domain from the opposite subunit at the rim of the active-site entryway, thereby leaving a large accessible area for substrates and interaction partners. When the structure of CsdA_{per} is compared with the structure of SufS (PDB 1JF9),¹⁹ it becomes apparent that several N-terminal motifs in SufS have been removed or reshaped in CsdA to increase access to the active site, including the deletion of the residues forming a β -hairpin in SufS (residues 18–22) and the bending of the long H3 helix at the dimer interface after a 45° turn at Glu66-Ala67.

Further oxidative modification of persulfated Cys358 by dissolved O_2 leads to the formation of an S-sulfation modification that cannot be repaired in bacteria, and which is unambiguously observed in the structure of $(\text{CsdA}_{\text{ss}})_2$ (Figure 2 B). The reactive nature of persulfated Cys358 is underscored by the fast conversion to the S-sulfated modification by exposure to air. The bulkier S-sulfate group

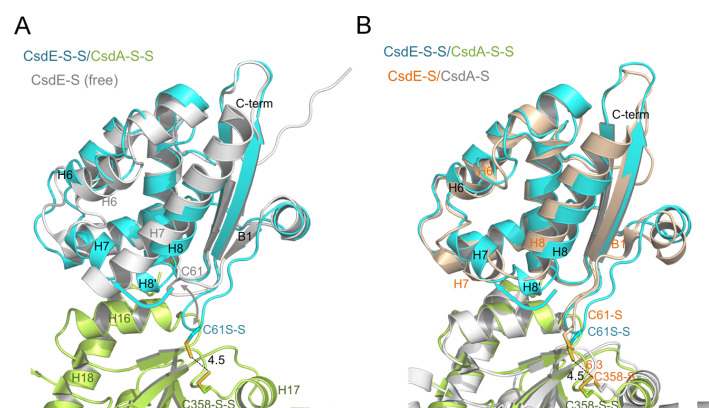


Figure 3. Conformational changes in bound CsdE_{per} facilitate transpersulfuration. (A) Superposition of the NMR structure of free CsdE (PDB 1NI7, in gray)³⁰ onto the crystal structure of the persulfurated complex (CsdA_{per} in green, CsdE_{per} in cyan). The superposition highlights the concerted structural rearrangements that lead to the newly formed H7 (in bound CsdE_{per}) and the splitting of helix H7 (in free CsdE) into H8' and H8 in bound CsdE_{per} (blue arrow). The orientation between the new H7, H8', and H8 is approximately 85° (blue arrow). These structural changes result in the swinging of Cys61 from a semiburied conformation in free CsdE to a fully extended conformation. This rearrangement allows the approach of the acceptor and donor Cys residues in the complex (gray arrow). (B) Superposition of the crystal structure of the persulfurated complex (CsdA_{per} in green, CsdE_{per} in cyan) with the heterodimer of the apo complex (PDB 4LW4; CsdA in white, CsdE in wheat)²⁰ with a traced Cys61. Helices H7–H8 of CsdE in the superimposed apo complex exhibit an intermediate conformation between free CsdE and bound CsdE_{per}. The distance between Cys61Sδ and Cys358Sδ in the persulfurated complex is shown in black dashed lines, while that between Cys61Sγ and Cys358Sγ in the apo complex is shown in orange.

is easily accommodated into the wide active site by the same residues (see below) that stabilize the persulfuration modification thereby impairing the S transfer process to any downstream acceptor (CSD or SUF system). The ease with which the S-sulfation of persulfurated Cys358 occurs by exposure to air agrees with the fact that a functional CSD system would be impaired during oxidative stress.⁴⁴

Crystal Structure of the Persulfurated (CsdA_{per}–CsdE_{per})₂ Complex. To gain insight into the structural basis and the mechanism by which persulfuration occurs across the interface between cysteine desulfurases and their sulfur acceptor cosubstrates, we determined the 2.49 Å resolution X-ray crystal structure of the persulfurated (CsdA_{per}–CsdE_{per})₂ complex from *Escherichia coli*. The asymmetric unit of the (CsdA_{per}–CsdE_{per})₂ crystal contained four independent heterotetramers with a solvent content of 54%. Three heterotetramers were well ordered in the crystal and served as the basis for our analyses; the fourth tetramer suffered from static disorder and was not considered further. The heterocomplex is formed by two subunits of CsdE_{per} bound to each active site of the CsdA_{per} homodimer, with the small helical domain burying a surface of 750 Å². The cosubstrate CsdE, the S acceptor in the CSD system, showed a fold consisting of a two-layered α/β sandwich with a core three-stranded β-sheet surrounded by a bundle of five α-helices (yellow and blue, Figure 1 A). The crystal structures of CsdA, CsdE, and the heterodimeric complex were in complete agreement with the corresponding structures in solution as judged by small-angle X-ray scattering data (SAXS, SI Text, and Figure S3).

The CsdA_{per}–CsdE_{per} interface comprises helices H16 and H18 from CsdA_{per} and helices H2 and H4 and the loop around the acceptor Cys61 from CsdE_{per} (Figure 3A). While CsdA remains essentially unchanged in the persulfurated complex, the structure of bound CsdE_{per} undergoes dramatic conformational rearrangements that affect the loop connecting helices H6–H7

and the α-helix H7 of the free form of CsdE. In bound CsdE_{per}, the H6–H7 loop becomes a α-helix (new H7) and the free conformation of the old H7 becomes distorted into two helical segments, H8' (a 3₁₀ helix) and H8, a shortened C-terminal α-helix (Figure 3A). These concerted structural changes help to maintain the spacing and the 85° tilt between helices H7 and H8 in the bound conformation of CsdE_{per} (Figure 3, blue arrow). Furthermore, this conformational change allows the assembly of the catalytically active complex by locating the acceptor residue, Cys61, to the proper position for the reaction to take place, closer to the donor Cys358 (11.8 Å between the Sγ atoms) with respect to the free form of CsdE (Figure 3A, gray arrow). In the (CsdA_{per}–CsdE_{per})₂ crystal structure the two persulfidic Sδ atoms, which can be unambiguously traced in the electron density maps (Figure 2 C), are separated merely by a ~4.5 Å gap.

The persulfurated (CsdA_{per}–CsdE_{per})₂ structure also underscores the dynamic nature of the CsdA–CsdE interaction. When six independent heterodimers are superimposed using only the CsdA coordinates, only the CsdE Cys61 tip residues adopt a fixed position and orientation, while the complete CsdE_{per} subunit swings about a hinge axis defined by CsdA_{per} α-helix H16 (residues 338–349) with an angular spread of 8° (Figure S4). This degree of structural plasticity in both CsdE and in the CsdA–CsdE interface may facilitate the efficient transfer of CsdA Cys358Sδ to CsdE Cys61Sγ. In fact, it allows the structural motif harboring the acceptor Cys61 side chain to be placed deep enough into the more rigid active site of persulfurated CsdA.

Solution CsdA–CsdE Interface Recapitulates the Persulfurated Interface. The mutual orientation and overall interactions between CsdA_{per} and CsdE_{per} observed in the crystal is also observed in the apo complex in solution by NMR techniques. Spectral perturbations were observed that allowed mapping the behavior of specific residues onto the

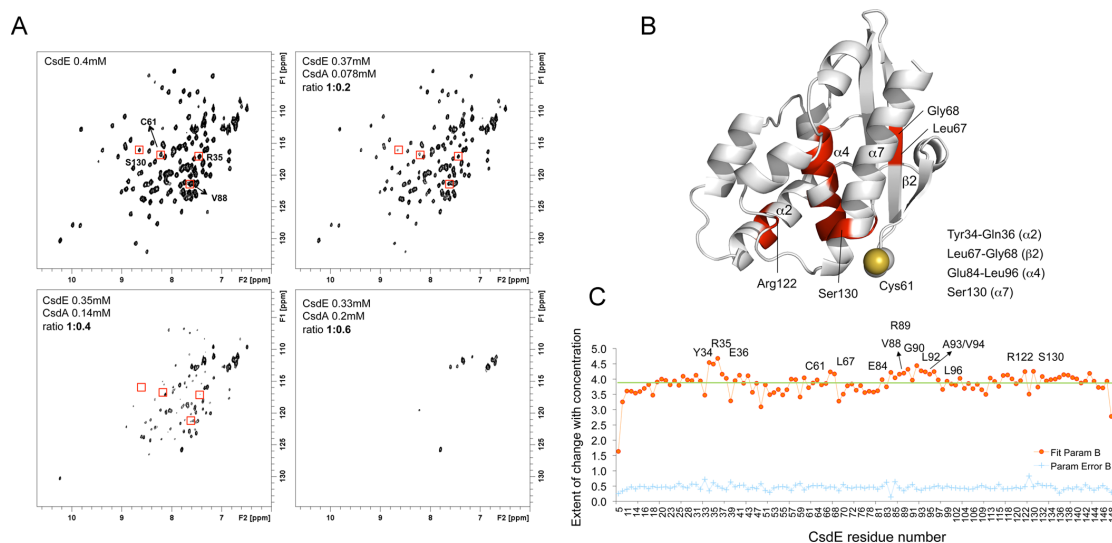


Figure 4. (CsdA–CsdE)₂ interface probed by NMR. (A) ^1H – ^{15}N HSQC spectra of 0.4 mM uniformly labeled ^{15}N -CsdE either free in solution or upon addition of unlabeled CsdA to 1:0.2, 1:0.4, and 1:0.6 molar ratios. (B) Ribbon representation of the solution NMR CsdE structure (PDB 1NI7). Amino acids whose NH backbone correlation crosspeak volume in the HSQC spectrum decreases to a larger extent than the average upon CsdA addition are colored in red and boxed by red squares in part A. Cys61 is indicated in spheres (S in yellow, C in white). (C) Plot of the decay rate of crosspeak volume (red circles) of CsdE residues upon CsdA addition, according to an exponential decay function $V = Ae^{-Bx}$, where V is the volume of an NH crosspeak, A is a scaling factor, and x is the CsdE:CsdA molar ratio. Blue crosses represent the fitting error on B. Residues with the fastest decay of the volume crosspeak (largest B) are annotated.

protein–protein interface. In particular, we used uniformly ^{15}N labeled CsdE that was titrated with unlabeled CsdA up to a 1:0.6 molar ratio. Most NH crosspeaks could be assigned and agreed with the previously published assignments.³⁰ ^1H – ^{15}N HSQC spectra acquired for every titration point (Figure 4A) revealed a sharp decrease in the integrated peak intensity for all the backbone NH CsdE correlations upon CsdA addition, while the intensity of the crosspeaks corresponding to the amide side chains were far less perturbed. These observations corroborate the formation of a high affinity (CsdA–CsdE)₂ complex in solution. Fitting the decreasing volume of the NH crosspeaks to an exponential decay function, we identified the CsdE residues mostly affected by CsdA binding: Tyr34, Arg35, Glu36, Cys61, Leu67, Glu84, Val88, Arg89, Gly90, Leu92, Ala93, Val94, Leu96, Arg122, and Ser130 (Figure 4B and C). These residues map onto the three α -helices of CsdE that lie at the interface with CsdA, therefore delineating the CsdA–CsdE solution interface (Figure 4B). Fittingly, this observation is in perfect agreement with the interface observed in the crystallographic structure.

Stabilization of the Transpersulfuration Reaction between CsdA and CsdE Requires the Canonical Sequence Motif of Group II Cys Desulfurases. The composite active site revealed by the crystal structure of persulfurated (CsdA_{per}–CsdE_{per})₂ has important consequences for the mechanism of transpersulfuration. As in the case of the intramolecular persulfuration of Cys358 by L-Cys, the mechanism for the interprotein sulfur transfer reaction can be proposed to consist in a nucleophilic attack by the deprotonated CsdE_{apo} Cys61S γ onto the persulfurated CsdA_{per} Cys358S δ . A preexisting CsdE_{apo} Cys61S γ thiolate can be assumed to take place. This hypothesis is based on the

experimentally determined pK_a for Cys61 of 6.5. This value is nearly two pH units lower than that of free cysteine and validates the presence of a predominantly ionized thiolate (>88%).⁴⁴ Likewise, the measured pK_a value for the equivalent Cys51 in the sequence and structurally homologous SufE is 6.3, therefore consistent with the parallel functional roles played by CsdE and SufE as recipients of persulfidic S atoms. The postulated nucleophilic attack requires the two sulfur atoms to be within a 4.5 Å distance, which appeared difficult to reconcile with the inter-S γ distances observed in the available crystal structure of the unmodified complex,²⁰ where no obvious interactions were found to stabilize a productive reaction intermediate. In the persulfurated (CsdA_{per}–CsdE_{per})₂ structure reported herein, however, a stabilizing chemical environment can be observed as a result of a productive CsdA–CsdE binding configuration. The optimal geometry for transpersulfuration requires the reorientation of the reactive Cys side chains to align Cys61 C β –S γ and Cys358 S γ –S δ bonds, in order to bring the thiol sulfur atoms into closer proximity. This configuration is compatible with the structure of the active site observed in the (CsdA_{per}–CsdE_{per})₂ complex (Figure 2C). The Cys358 persulfide of CsdA is stabilized by His120N δ , Lys250* from the other subunit, and Cys358's own backbone amide atoms through a water-mediated contact (Figure 2A,C). As expected, these interactions with Cys358 S δ –S γ are also observed in the functional persulfurated (CsdA_{per})₂ active site in the presence of CsdE (Figure 2A). In the persulfurated complex, CsdA_{per} Cys358 loop residues stabilize CsdE_{per} Cys61S γ and S δ atoms via direct and water-mediated hydrogen bonds with the backbone NH groups of Gln356, His357, and Cys358 (Figure 2C). Glu62 is the only other residue from the Cys61 hairpin that interacts with CsdA, specifically through the Glu62

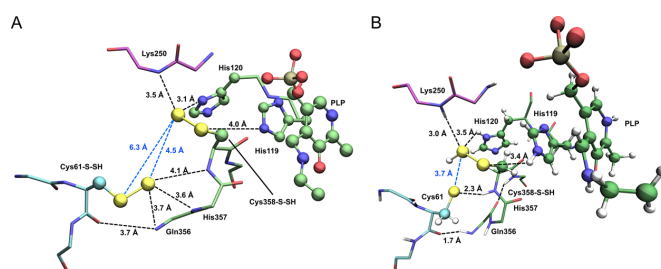


Figure 5. Catalytic residues involved in transsulfuration across the $(\text{CsdA}-\text{CsdE})_2$ interface. (A) Persulfurated $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$ crystal structure, with both catalytic Cys residues as persulfides. (B) Insight into the active site from the minimized structure obtained at the M06-2X/6-31G**/MM level of theory, with Cys61 modeled with a thiolate side chain. The atoms belonging to the QM subsystem are shown as ball and sticks. Carbon atoms are depicted in cyan (CsdE) and green (CsdA_{per}) in (A) and dark olive green in (B), with relevant interactions shown as dashed lines with their interatomic distances shown in angstroms (inter-S interactions are shown in blue).

carboxylate group and Glu62O to the carbonyl Ala354O and to His338Nε2 respectively. This hydrogen bond network provides the chemical rationale for the conservation of the amino acid sequence in the Cys358 loop of CsdA (³⁵⁶QHCAQ³⁶⁰), and, conceivably, it can be extended to the other class II cysteine desulfurases where this sequence signature is conserved, including SufS.¹³ The fully oxidized state of persulfurated Cys358 Sγ-Sδ impairs the reactivity of the sulfur atom by blocking the S transfer (Figure 2B).

Residues from CsdA Cys-Loop Further Assist Catalysis by Stabilizing the Nucleophilic CsdE Cys61 Thiolate. The CsdA_{per}-CsdE_{per} interface creates the optimum environment for the relocation of Cys61 Sγ in a favorable orientation for transsulfuration. To shed light into the mechanism of the transsulfuration reaction, we resorted to computational chemistry calculations using the persulfurated $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$ complex structure as the template and removing Cys61 Sδ from CsdE (Figure 5). We performed molecular dynamics (MD) simulations with the AMBER 99SB force field for the $(\text{CsdA}_{\text{per}}-\text{CsdE})_2$ complex where the Cys61 was modeled as a thiolate (see the SI Text and Figures S5 and S6 for details). The total simulation time was 50 ns. Afterward, the system was minimized by means of the hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) methodology at the M06-2X/6-31G**/MM level of theory, where the QM subsystem was composed of the side chains of Cys61, Cys358, His119, His120, and the PLP cofactor (69 atoms in total). The resulting structure for the active site (Table S3) is depicted in Figure 5B.

Analysis of the MD simulation showed that the CsdE subunit displays a significant flexibility, also in agreement with the previous observations. The Sδ-Sγ distance fluctuates between 3.5 and 8.5 Å (Figure 6), a distance variation that can be accounted for by the inherent pivoting of CsdE over the CsdA active site groove, as also observed in the crystal structure of the persulfurated complex (Figure S4). Thus, during the 50 ns trajectory the system visited conformations where the direct transfer from CsdA_{per} Cys358 Sδ to CsdE_{apo} Cys61 Sγ is feasible. On the contrary, equivalent MD simulations of the system with a protonated CsdE Cys61 showed a substantially larger Sδ-Sγ distance, for which the direct transsulfuration is not possible (Figure S6). This result confirms that Cys61 must be deprotonated in order to favor the process.

The QM/MM optimized $(\text{CsdA}_{\text{per}}-\text{CsdE})_2$ complex exhibited a reaction-competent geometry for the thiol-persulfido exchange reaction between CsdA_{per} Cys358 Sδ and CsdE_{apo}

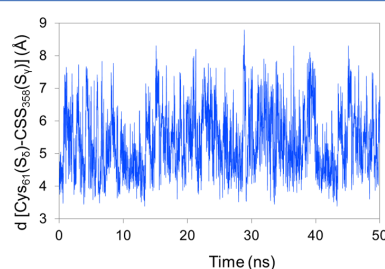


Figure 6. Intersulfur distance fluctuations during MD simulations. Fluctuations of the distance (Å) between the Sδ atom of CsdE Cys61 and the Sγ atom of the persulfurated CsdA Cys358 observed during 50 ns of MD simulation. CsdE Cys61 Sδ was modeled in the ionized form (thiolate), as this is the physiologically relevant form (experimental $\text{pK}_a = 6.5$).

Cys61 Sγ. The Cys61 Sγ atom is found at 3.7 Å from the persulfurated Cys358 Sδ, thereby suggesting the possibility of a direct nucleophilic attack, which could lead to the persulfuration of CsdE. The position and orientation of CsdE Cys61 is favored by the hydrogen bond network established with the backbone NH groups of CsdA Gln356 and Cys358. In turn, persulfurated Cys358 from CsdA is kept by means of hydrogen bonds to the imidazole rings of His119 and His120 and the backbone amide group of Lys250*. The positive polarization generated by the Cys358-loop backbone structure, specifically through the electrostatic stabilization afforded by the aligned amide NH functions of Gln356 and Cys358 from the Cys loop, would further assist catalysis by stabilizing the Cys61 thiolate over the corresponding thiol and/or the negative charge spread over the transition state.

To gain further insight into the enzymatic mechanism of the transsulfuration, we explored the potential energy surface (PES) along the antisymmetric Sδ-Sγ-Sδ stretch coordinate, which delineates the pathway for interprotein S transfer. A transition state was found, with a potential energy barrier of 20.2 kcal mol⁻¹. The corresponding geometry is presented in Figure 7, showing that the protonated Sγ is placed midway between the Sδ atoms of CsdE Cys61 (2.4 Å) and CsdA Cys358 (2.9 Å). According to this structure, the nucleophilic attack of the Cys61 thiolate on the persulfurated Cys358 is favored by the hydrogen bond interactions of the side chains of His119 and Lys250* with the Sγ atom of Cys358, as well as by the electrostatic stabilization provided by PLP. In this scenario,

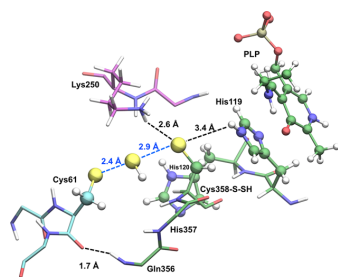


Figure 7. Transition state geometry for the transpersulfuration reaction. Ball-and-stick representation of the transition state geometry observed during MD simulations of the $(\text{CsdA}_{\text{per}}-\text{CsdE})_2$ structure. Color codes for carbon atoms are cyan (CsdA) and green and violet (CsdE subunits). The S_γ atom is shown on the path between the donor CsdA Cys358 S_δ (modeled initially as a thiol function) and the acceptor CsdE Cys61 S_δ (modeled initially as a thiolate), with intersulfur interactions depicted as dashed lines with distance labels in blue, and all other interactions and distances in black. It is noteworthy that the key participation of His119 and Lys250* (* denotes the other subunit) for stabilizing the negative charge that develops on CsdA Cys358 S_δ as the transpersulfuration progresses and ultimately in its reprotonation process.

the CsdA Cys358 thiolate formally developed after the direct S transfer to the acceptor CsdE Cys61 would be reprotonated by His119 and/or Lys250*. One of these two residues, Lys250 could easily reprotonate the temporary thiolate generated on Cys358 δ by taking up protons from the solvent, since it is highly exposed to the bulk solvent.

The role of His119, His120, and Lys250* in mediating the transpersulfuration could be extrapolated to other sulfur donor–acceptor pairs since they are either conserved in SufS (His123, His124) or substituted by a potentially isofunctional group (His55). In IscS , the equivalent residues can be identified from the superposition of the CsdA and IscS in complex with their S acceptors; His343 in IscS could take the place of His119 and/or His120 in providing stabilization to the catalytic Cys321, while there is no clear counterpart to Lys250* in CsdA except for IscS Lys99, whose terminal ammonium side chain could potentially substitute for Lys250* amide main-chain function. The conservation of functional roles across cysteine desulfurases suggests that the fundamental chemical reaction underlying the transfer of S atoms across protein–protein interfaces could be essentially conserved across desulfurases in spite of considerable sequence divergence. Furthermore, the sequence variation at His119, His120, and Lys250* positions could be partly responsible for the greater efficiency of CsdA as a cysteine desulfurase enzyme when compared with SufS/ IscS .⁹ If true, the conservation of the active site structure (Cys loop) plus the tunability of additional catalytic residues (e.g., His119, Lys250*) would underscore the conservation of the mechanistic processes of the transpersulfuration reaction.

CONCLUSION

In conclusion, the sequence and structural similarities between CsdA / CsdE and SufS/SufE suggest the existence of a common sulfur acceptor–donor recognition mechanism shared by the CSD and SUF systems. The conformational changes and intersubunit adjustments observed in the persulfurated $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$ complex result in a tighter and more

extensive interface than in a previous report²⁰ (SI text). In the persulfurated $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$ interface, the distance between the reactive persulfide and thiol groups (CsdA_{per} Cys358 S_δ and CsdE_{per} Cys61 S_γ) becomes 0.5–1 Å shorter than in the unmodified complex, reaching a more favorable position and orientation for transpersulfuration (Figures 1B and 5A). Since in the $(\text{CsdA}-\text{CsdE})_2$ complexes (persulfurated and unmodified) the access to the catalytic center is unimpeded, CsdA can achieve an essentially continuous persulfurated state, whereby transpersulfuration of CsdE would be closely followed by the repersulfuration of CsdA Cys358. In keeping with the modest effect of CsdE in CsdA desulfurase activity (a 2-fold increase⁹), the role of CsdE in enhancing CsdA activity could be limited to provide an acceptor Cys for Cys358 persulfide. The MD simulations and QM/MM calculations and the analysis of the active site in the crystal structures of persulfurated $(\text{CsdA}_{\text{per}})_2$ and $(\text{CsdA}_{\text{per}}-\text{CsdE}_{\text{per}})_2$ demonstrate the existence of an electrostatic stabilization role for the Cys358 loop of CsdA (QHCAQ), the signature motif in SufS-like group II cysteine desulfurases that includes CsdA and SufS, and suggest stabilizing/catalytic roles for the conserved His119, His120, and the main chain and side chain of the less conserved Lys250* residue, in particular for stabilizing the developing negative charge on Cys358 S_δ and its subsequent reprotonation. The conservation of these critical residues enables the occasional crosstalk between systems without compromising their specificity. In group II Cys desulfurases in particular, the relative specificity of $\text{CsdA}-\text{CsdE}$ and SufS–SufE pairs is generated by evolutionary divergence of secondary structure motifs (i.e., the elongation of strand β_2 and the C-terminal helix on the surface of CsdA) that leads to the optimization of the docking surface area of the Cys desulfurase for each system. In summary, group II cysteine desulfurases (and, possibly, group I cysteine desulfurases like IscS) have evolved toward versatile interaction surfaces that provide discrimination between their specific acceptors while deploying the precise chemical environment for the transpersulfuration reaction that mobilizes sulfane sulfur following a shared mechanism based in a highly conserved peptide motif near the catalytic Cys residues.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.6b00360.

Experimental procedures, X-ray diffraction data processing, structure refinement tables, ITC, NMR, SAXS, more detailed comparisons of the apo and persulfurated $(\text{CsdA}-\text{CsdE})_2$ complex, and details of the computational chemistry calculations and the optimized geometry of the QM optimized subsystem (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: cvega@cib.csic.es, cristina.vega@strubicib.org (M.C.V.).

Present Address

•(A.A. and J.J.-B.) Structural Biology, CIC bioGUNE, Derio-Bizkaia, Spain.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was funded by Spanish institutional grants BFU2008-02372/BMC, CSD 2006-23, and BFU2011-22588 to M.C., CTQ2012-36253-C03-03 and CTQ2015-66223-C2 to I.T., CTQ2015-64597-C2-1-P to J.J.B., and BFU2010-22266-C02-02 and CTQ2015-66206-C2-2-R to M.C.V. Further support for this work was obtained from the Generalitat Valenciana (ACOMP/2015/239 to I.T.) and from the European Commission FP7 ComplexINC grant (contract no. 279039) to M.C.V. We acknowledge the ESRF for provision of synchrotron radiation at ID23-1, ID23-2, ID14-1, and BM29 BioSAXS beamlines. E.P.S. and M.L.E. acknowledge the support of the Ph.D. program in Structure and Function of Proteins of the *Universitat Autònoma de Barcelona* (UAB) and the Ph.D. program in Biochemistry, Molecular Biology and Biomedicine of the *Universidad Complutense de Madrid* (UCM), respectively.

■ REFERENCES

- (1) Kessler, D. *FEMS Microbiol. Rev.* **2006**, *30*, 825–840.
- (2) Mueller, E. G. *Nat. Chem. Biol.* **2006**, *2*, 185–194.
- (3) Begley, T. P.; Xi, J.; Kinsland, C.; Taylor, S.; McLafferty, F. *Curr. Opin. Chem. Biol.* **1999**, *3*, 623–629.
- (4) Marquet, A. *Curr. Opin. Chem. Biol.* **2001**, *5*, 541–549.
- (5) Ayala-Castro, C.; Saini, A.; Outten, F. W. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 110–25 table of contents.
- (6) Fontecave, M.; Ollagnier-de Choudens, S. *Arch. Biochem. Biophys.* **2008**, *474*, 226–237.
- (7) Johnson, D. C.; Dean, D. R.; Smith, A. D.; Johnson, M. K. *Annu. Rev. Biochem.* **2005**, *74*, 247–281.
- (8) Fontecave, M.; Choudens, S. O.; Py, B.; Barras, F. *JBIC, J. Biol. Inorg. Chem.* **2005**, *10*, 713–721.
- (9) Loiseau, L.; Ollagnier-de Choudens, S.; Lascoux, D.; Forest, E.; Fontecave, M.; Barras, F. *J. Biol. Chem.* **2005**, *280*, 26760–26769.
- (10) Trotter, V.; Vinella, D.; Loiseau, L.; Ollagnier de Choudens, S.; Fontecave, M.; Barras, F. *Mol. Microbiol.* **2009**, *74*, 1527–1542.
- (11) Miyauchi, K.; Kimura, S.; Suzuki, T. *Nat. Chem. Biol.* **2012**, *9*, 105–111.
- (12) López-Esteva, M.; Ardá, A.; Savko, M.; Round, A.; Shepard, W. E.; Bruix, M.; Coll, M.; Fernández, F. J.; Jiménez-Barbero, J.; Vega, M. C. *PLoS One* **2015**, *10*, e0118606.
- (13) Mihara, H.; Esaki, N. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 12–23.
- (14) Zheng, L.; Cash, V. L.; Flint, D. H.; Dean, D. R. *J. Biol. Chem.* **1998**, *273*, 13264–13272.
- (15) Mihara, H.; Maeda, M.; Fujii, T.; Kurihara, T.; Hata, Y.; Esaki, N. *J. Biol. Chem.* **1999**, *274*, 14768–14772.
- (16) Pan, J.; Carroll, K. S. *ACS Chem. Biol.* **2013**, *8*, 1110–1116.
- (17) Shi, R.; Proteau, A.; Villarroya, M.; Moukadiri, I.; Zhang, L.; Trempe, J. F.; Matte, A.; Armengod, M. E.; Cygler, M. *PLoS Biol.* **2010**, *8*, e1000354.
- (18) Marinoni, E. N.; de Oliveira, J. S.; Nicolet, Y.; Raulfs, E. C.; Amara, P.; Dean, D. R.; Fontecilla-Camps, J. C. *Angew. Chem., Int. Ed.* **2012**, *51*, 5439–5442.
- (19) Lima, C. D. *J. Mol. Biol.* **2002**, *315*, 1199–1208.
- (20) Kim, S.; Park, S. *J. Biol. Chem.* **2013**, *288*, 27172–27180.
- (21) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125–132.
- (22) Evans, P. R. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 282–292.
- (23) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.
- (24) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352–367.
- (25) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486–501.
- (26) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12–21.
- (27) Vranken, W. F.; Boucher, W.; Stevens, T. J.; Fogh, R. H.; Pajon, A.; Llinas, M.; Ulrich, E. L.; Markley, J. L.; Ionides, J.; Laue, E. D. *Proteins: Struct., Funct., Genet.* **2005**, *59*, 687–696.
- (28) Grzesiek, S.; Bax, A. *J. Magn. Reson. (1969-1992)* **1992**, *99*, 201–207.
- (29) Grzesiek, S.; Bax, A. *J. Am. Chem. Soc.* **1992**, *114*, 6291–6293.
- (30) Liu, G.; Li, Z.; Chiang, Y.; Acton, T.; Montelione, G. T.; Murray, D.; Szyperski, T. *Protein Sci.* **2005**, *14*, 1597–1608.
- (31) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Luchko, T.; Luo, R.; Mandej, B.; Merz, K. M. *The Amber Molecular Dynamics Package*; University of California: San Francisco, 2015.
- (32) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (33) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
- (34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (35) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (36) Aranda, J.; Zinovjev, K.; Roca, M.; Tuñón, I. *J. Am. Chem. Soc.* **2014**, *136*, 16227–16239.
- (37) Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K. *J. Phys. Chem. B* **2015**, *119*, 917–927.
- (38) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088.
- (39) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.
- (40) Prisch, F.; Konarev, P. V.; Iannuzzi, C.; Pastore, C.; Adinolfi, S.; Martin, S. R.; Svergun, D. I.; Pastore, A. *Nat. Commun.* **2010**, *1*, 95.
- (41) Singh, H.; Dai, Y.; Outten, F. W.; Busenlehner, L. S. *J. Biol. Chem.* **2013**, *288*, 36189–36200.
- (42) Selbach, B.; Earles, E.; Dos Santos, P. C. *Biochemistry* **2010**, *49*, 8794–8802.
- (43) Urbina, H. D.; Silberg, J. J.; Hoff, K. G.; Vickery, L. E. *J. Biol. Chem.* **2001**, *276*, 44521–44526.
- (44) Bolstad, H. M.; Botelho, D. J.; Wood, M. J. *J. Proteome Res.* **2010**, *9*, 5358–5369.

